BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: April 19, 2003, 05:43:38 ; Search time 547 Seconds
(without alignments)
9658.479 Million cell updates/sec

Title: US-09-920-677-3
Perfect score: 2346
Sequence: 1 gcacgaggctgcggcgggtc.....gtaacccagctgcggagcct 2346

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2274872

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	60	2.6	60	24	ABN41230	Human spliced tran
c	2	51	2.2	51	22	AAL28838	Human SNP oligonuc
c	3	50	2.1	50	22	AAL28841	Human SNP oligonuc
c	4	40	1.7	51	22	AAL28839	Human SNP oligonuc
c	5	40	1.7	51	22	AAL28840	Human SNP oligonuc
c	6	26.2	1.1	51	22	AAL29715	Human SNP oligonuc
	7	25.6	1.1	51	22	AAH79450	Human DNA containi
c	8	25.2	1.1	51	22	AAL30783	Human SNP oligonuc
	9	25	1.1	25	23	AAS19268	Human protein kina
c	10	25	1.1	25	23	AAS19269	Human protein kina
	11	24.4	1.0	60	24	ABN59017	Human spliced tran
	12	24.2	1.0	57	18	AAV76048	Staphylococcus aur
	13	24	1.0	30	17	AAT32212	Rat mitogen-activa
	14	23.8	1.0	32	17	AAT32213	Yeast ACE1-rat mit
c	15	23.8	1.0	49	22	AAL29254	Human SNP oligonuc
c	16	23.8	1.0	60	24	ABN45373	Human spliced tran
	17	23.8	1.0	60	24	ABN50683	Human spliced tran
	18	23.6	1.0	50	23	ABL00050	Human silent nonco
c	19	23.6	1.0	51	22	AAL28430	Human SNP oligonuc
	20	23.6	1.0	60	22	AAH84250	Human cell death p
	21	23.4	1.0	58	24	AAS19340	Oligonucleotide 59
c	22	23.2	1.0	50	16	AAT22872	Human gene signatu
c	23	23.2	1.0	50	22	AAL28930	Human SNP oligonuc
c	24	23.2	1.0	51	22	AAL31877	Human SNP oligonuc
	25	23.2	1.0	60	24	ABN34677	Human spliced tran
c	26	22.8	1.0	41	22	AAH44691	Human type-I amino
c	27	22.8	1.0	47	21	AAZ67514	Human map-related
	28	22.8	1.0	48	20	AAZ23542	Human DNA fragment
c	29	22.8	1.0	48	24	ABN71559	Streptococcus agal
	30	22.8	1.0	50	22	AAL28956	Human SNP oligonuc
	31	22.8	1.0	51	22	AAL32077	Human SNP oligonuc
c	32	22.8	1.0	60	21	AAC78364	Human cancer assoc
c	33	22.6	1.0	47	21	AAZ66546	Human map-related
c	34	22.6	1.0	47	21	AAZ69126	Human map-related
c	35	22.6	1.0	51	22	AAH38300	Human SNP flanking
	36	22.6	1.0	54	19	AAV65911	HCMV target DNA fo
	37	22.6	1.0	60	24	ABN35460	Human spliced tran
c	38	22.6	1.0	60	24	ABN46663	Human spliced tran
	39	22.4	1.0	44	15	AAQ69313	Human sodium/potas
	40	22.4	1.0	44	18	AAT63775	Human Na/K ATPase
	41	22.4	1.0	44	20	AAX17063	Test sequence from
	42	22.4	1.0	44	24	ABK82554	DNA binding molecu
c	43	22.4	1.0	48	22	AAF29313	Primer base sequen
c	44	22.4	1.0	48	22	AAF29318	Primer base sequen
c	45	22.4	1.0	51	22	AAL27843	Human SNP oligonuc
	46	22.4	1.0	51	22	AAL32091	Human SNP oligonuc
	47	22.2	0.9	45	15	AAQ69462	Human interleukin
	48	22.2	0.9	45	18	AAT63924	Human interleukin-
	49	22.2	0.9	45	20	AAX17212	Test sequence from
	50	22.2	0.9	45	24	ABK82703	DNA binding molecu
	51	22.2	0.9	48	22	AAF29294	Primer base sequen
	52	22.2	0.9	51	22	AAL32145	Human SNP oligonuc
	53	22.2	0.9	60	24	ABN40059	Human spliced tran
c	54	22.2	0.9	60	24	ABN45158	Human spliced tran
c	55	22	0.9	33	22	AAH44211	p70 S6K mutagenesi
c	56	22	0.9	39	24	ABA96305	LDL receptor allel
	57	22	0.9	42	20	AAX76470	Human BRCA1 intera
	58	22	0.9	47	21	AAZ67996	Human map-related
	59	22	0.9	47	21	AAA64611	AU rich sequence i
	60	22	0.9	48	22	AAF29302	Primer base sequen
c	61	22	0.9	50	23	ABL00059	Human silent nonco
c	62	22	0.9	51	22	AAL30100	Human SNP oligonuc
c	63	22	0.9	51	22	AAL31018	Human SNP oligonuc
c	64	22	0.9	51	22	AAL31438	Human SNP oligonuc
	65	22	0.9	51	22	AAH26598	Human GM-CSF gene
c	66	22	0.9	58	24	AAS19341	Oligonucleotide 59
	67	22	0.9	60	19	AAV04153	Blood group antige
	68	22	0.9	60	22	AAH93339	Plasmodium falcipa
	69	22	0.9	60	24	ABN38032	Human spliced tran
c	70	21.8	0.9	50	18	AAT61581	VH and scFv antibo
c	71	21.8	0.9	50	18	AAT61565	scFv antibody libr
	72	21.8	0.9	51	22	AAL28954	Human SNP oligonuc
	73	21.8	0.9	51	22	AAI76866	Human silent SNP c
	74	21.8	0.9	55	18	AAV76001	Staphylococcus aur
c	75	21.8	0.9	60	22	AAF56143	Staphylococcus aur
	76	21.8	0.9	60	24	ABN35167	Human spliced tran
c	77	21.8	0.9	60	24	ABN43764	Human spliced tran
	78	21.8	0.9	60	24	ABN50818	Human spliced tran
c	79	21.6	0.9	51	22	AAL30101	Human SNP oligonuc
c	80	21.6	0.9	51	22	AAL30786	Human SNP oligonuc
c	81	21.6	0.9	51	22	AAI76504	Human silent SNP c
c	82	21.6	0.9	51	22	AAI76505	Human silent SNP c

c	83	21.6	0.9	54	16	AAT23964	Human gene signatu
	84	21.6	0.9	60	24	ABN36856	Human spliced tran
	85	21.4	0.9	48	19	AAV69937	Chlorella virus pr
	86	21.4	0.9	48	24	AAS21106	(GGA)16 DNA purifi
	87	21.4	0.9	48	24	AAD24852	Chlorella virus cv
	88	21.4	0.9	51	22	AAL31468	Human SNP oligonuc
c	89	21.4	0.9	51	22	AAI78865	Human silent SNP c
c	90	21.4	0.9	53	22	AAH93331	Plasmodium falcipa
	91	21.4	0.9	59	21	AAC11350	Human secreted pro
c	92	21.4	0.9	60	24	ABN32960	Human spliced tran
	93	21.4	0.9	60	24	ABN37046	Human spliced tran
c	94	21.4	0.9	60	24	ABN37096	Human spliced tran
c	95	21.4	0.9	60	24	ABN50863	Human spliced tran
c	96	21.2	0.9	41	22	AAH44690	Human type-I amino
c	97	21.2	0.9	48	22	AAF29312	Primer base sequen
c	98	21.2	0.9	50	22	AAI77361	Human silent SNP c
c	99	21.2	0.9	51	21	AAA77364	Human clone cg4491
c	100	21.2	0.9	51	22	AAL31990	Human SNP oligonuc

ALIGNMENTS

RESULT 1
ABN41230
ID ABN41230 standard; DNA; 60 BP.
XX
AC ABN41230;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:13978.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 13978; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 20 A; 12 C; 18 G; 10 T; 0 other;

Query Match 2.6%; Score 60; DB 24; Length 60;
Best Local Similarity 100.0%; Pred. No. 9.4e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1468 ATAGAGCAGATGGATGTGACAATGAGTGGGGAAGCATCGGCACCACTTCCAATACGACAG 1527
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 ATAGAGCAGATGGATGTGACAATGAGTGGGGAAGCATCGGCACCACTTCCAATACGACAG 60

RESULT 2
AAL28838/c
ID AAL28838 standard; DNA; 51 BP.
XX
AC AAL28838;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #2046.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 1966; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of

CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 51 BP; 12 A; 9 C; 9 G; 21 T; 0 other;

Query Match 2.2%; Score 51; DB 22; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1707 TCAATGTCATTACATAGAACACTTCGGACACAGGAAAAATAAACGTGGATT 1757
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 51 TCAATGTCATTACATAGAACACTTCGGACACAGGAAAAATAAACGTGGATT 1

RESULT 3
AAL28841/c
ID AAL28841 standard; DNA; 50 BP.
XX
AC AAL28841;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #2049.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 1967; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 50 BP; 13 A; 17 C; 3 G; 17 T; 0 other;

Query Match 2.1%; Score 50; DB 22; Length 50;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1610 GCAATGCTTTTAATGAATTTAAGGCAAAAAGGTGGAGAGGGAGATGTGTG 1659
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 50 GCAATGCTTTTAATGAATTTAAGGCAAAAAGGTGGAGAGGGAGATGTGTG 1

RESULT 4
AAL28839/c
ID AAL28839 standard; DNA; 51 BP.
XX
AC AAL28839;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #2047.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 1966; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,

```
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 10 A; 14 C; 9 G; 18 T; 0 other;

  Query Match             1.7%;  Score 40;  DB 22;  Length 51;
  Best Local Similarity   98.1%;  Pred. No. 1.3;
  Matches   51;  Conservative    0;  Mismatches.   0;  Indels    1;  Gaps    1;

Qy   1655 GTGTGAGCATCCTGCAAGGTGAAACAAGACTCAAAATGACAGTTTCAGAGAG 1706
           ||||||||||||||||||||||| |||||||||||||||||||||||||||
Db     51 GTGTGAGCATCCTGCAAGGTGAAAC-AGACTCAAAATGACAGTTTCAGAGAG 1


RESULT 5
AAL28840/c
ID   AAL28840 standard; DNA; 51 BP.
XX
AC   AAL28840;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #2048.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 1967; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 10 A; 14 C; 9 G; 18 T; 0 other;

  Query Match             1.7%;  Score 40;  DB 22;  Length 51;
  Best Local Similarity   98.1%;  Pred. No. 1.3;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy   1655 GTGTGAGCATCCTGCAAGGTGAAACAAGACTCAAAATGACAGTTTCAGAGAG 1706
           ||||||||||||||||||||||| |||||||||||||||||||||||||||
Db     51 GTGTGAGCATCCTGCAAGGTGAAAC-AGACTCAAAATGACAGTTTCAGAGAG 1


RESULT 6
AAL29715/c
ID   AAL29715 standard; DNA; 51 BP.
XX
AC   AAL29715;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #2923.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 2223; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 8 A; 7 C; 3 G; 33 T; 0 other;

  Query Match             1.1%;  Score 26.2;  DB 22;  Length 51;
```

```
  Best Local Similarity   72.3%;  Pred. No. 4.8e+03;
  Matches   34;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   1741 AAAAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAA 1787
          ||| |||  | ||||  | || |||  ||  ||||||| ||||||||
Db     47 AAACATAGTCCTGGAAATAAATAAAAAAAAAAAATGGTGAAAAAAAAA 1

RESULT 7
AAH79450
ID   AAH79450 standard; DNA; 51 BP.
XX
AC   AAH79450;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human DNA containing single nucleotide polymorphism SEQ ID NO. 65.
XX
KW   Human; single nucleotide polymorphism; SNP; angiopoietin;
KW   4-hydroxybutyrate; dehydrogenase; protein therapy;
KW   adenosine triphosphate-dependent RNA helicase;
KW   major histocompatibility complex Class I histocompatibility antigen; MHC;
KW   phosphoglycerate kinase; immunosuppressive; immunostimulatory;
KW   antirheumatic; antisclerotic; antidiabetic; antiinflammatory; cytostatic;
KW   antileukemic; neuroprotective; antimicrobial; gene therapy; vaccine; ds.
XX
OS   Homo sapiens.
XX
PN   WO200148245-A2.
XX
PD   05-JUL-2001.
XX
PF   27-DEC-2000; 2000WO-US35346.
XX
PR   27-DEC-1999;   99US-0472688.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-418297/44.
XX
PT   Polymorphic nucleic acids encoding e.g. angiopoietin, dehydrogenase,
PT   adenosine triphosphate-dependent RNA helicase and/or phosphoglycerate
PT   kinase, useful for diagnosing and treating, e.g. cancer, autoimmune
PT   diseases and infections -
XX
PS   Claim 1; Page 77; 484pp; English.
XX
CC   The invention relates to nucleic acids (AAH79386-AAH80036) encoding
CC   polymorphic variants of proteins (AAG98010-AAG98238) related to
CC   angiopoietin, 4-hydroxybutyrate, dehydrogenase, adenosine triphosphate
CC   (ATP)-dependent RNA helicase, major histocompatibility complex (MHC)
CC   Class I histocompatibility antigen and/or phosphoglycerate kinase. These
CC   nucleic acid single nucleotide polymorphisms (SNPs) and the encoded
CC   proteins have potential immunosuppressive, immunostimulatory,
CC   antirheumatic, antisclerotic, antidiabetic, antiinflammatory, cytostatic,
CC   antileukemic, neuroprotective and antimicrobial activity and may be
CC   useful in gene/protein therapy, vaccines, modulation of the expression
CC   and activity of proteins related to angiopoietin, 4-hydroxybutyrate,
CC   dehydrogenase, adenosine triphosphate (ATP)-dependent RNA helicase,
CC   major histocompatibility complex (MHC) Class I histocompatibility antigen
CC   and/or phosphoglycerate kinase. Disorders that may be prevented,
CC   diagnosed and/or treated by the above methods include multifactorial
CC   diseases with a genetic component, such as autoimmune diseases (e.g.
CC   rheumatoid arthritis, multiple sclerosis, diabetes, systemic lupus
CC   erythromatosus and Grave's disease), inflammation, cancer (e.g. cancers
CC   of the bladder, brain, breast, colon and kidney, leukemia), diseases of
CC   the nervous system, an infection of pathogenic organisms. They may also
CC   be used to alter phenotypic traits such as longevity, appearance,
CC   strength, speed and endurance.
XX
SQ   Sequence 51 BP; 15 A; 11 C; 6 G; 19 T; 0 other;

  Query Match             1.1%;  Score 25.6;  DB 22;  Length 51;
  Best Local Similarity   70.8%;  Pred. No. 6.8e+03;
  Matches   34;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    665 GGATCATCTACAGAGACCTGAAGCCGGAGAATATCATGCTTAATCACC 712
          | |.|||| || | || || ||||| || |||||| | |||  | |||
Db      4 GTATCATTCACTGTGATCTAAAGCCTGAAAATATCCTTCTTTGTAACC 51

RESULT 8
AAL30783/c
ID   AAL30783 standard; DNA; 51 BP.
XX
AC   AAL30783;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #3991.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 2533; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 6 A; 4 C; 3 G; 38 T; 0 other;

  Query Match             1.1%;  Score 25.2;  DB 22;  Length 51;
  Best Local Similarity   71.7%;  Pred. No. 8.6e+03;
```

```
  Matches   33;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   1757 TTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
          ||| |||||| || || ||  |||| ||||||   |||  |||| |
Db     51 TTTCAAAAATAAAGCAGTGAAGCAATAAAAAAAAAAAAAAAAAAAA 6


RESULT 9
AAS19268
ID   AAS19268 standard; DNA; 25 BP.
XX
AC   AAS19268;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human protein kinases SGK351 PCR primer #1.
XX
KW   Human; ss; SGK351; protein kinase; cancers; immune-related disease;
KW   cardiovascular disease; neuronal-associated disease; metabolic disorder;
KW   tissue cancer; haematopoietic cancer; central nervous system disorder;
KW   peripheral nervous system disorder; Alzheimer's disease;
KW   Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW   viral infection; prion infection; bacterial infection; fungal infection;
KW   ocular disease; migraine; pain; sexual dysfunction; mood disorder;
KW   attention disorder; cognition disorder; hypotension; hypertension;
KW   psychotic disorder; neurological disorder; dyskinesia; gene therapy;
KW   metabolic disorder; organ transplant rejection; PCR primer.
XX
OS   Homo sapiens.
XX
PN   WO200177338-A2.
XX
PD   18-OCT-2001.
XX
PF   10-APR-2001; 2001WO-US11675.
XX
PR   10-APR-2000; 2000US-195953P.
PR   01-MAY-2000; 2000US-201015P.
PR   22-JUN-2000; 2000US-213805P.
XX
PA   (SUGE-) SUGEN INC.
XX
PI   Plowman G,  Whyte D,  Manning G,  Sudarsanam S,  Martinez R;
PI   Caenepeel S;
XX
DR   WPI; 2001-657173/75.
XX
PT   Novel kinase polypeptides useful for identifying modulators of kinase
PT   polypeptide that are used for inhibiting kinase activity for treating
PT   cancers, Alzheimer's disease, Parkinson's disease, multiple sclerosis
PT   -
XX
PS   Example 2b; Page 104; 167pp; English.
XX
CC   The invention relates to an isolated, enriched, or purified kinase
CC   polypeptide comprising a 1360 or 198 residue amino acid sequence, fully
CC   defined in specification, or comprising a kinase except which
CC   lacks one or more of N- or C-terminal domain, C-terminal catalytic
CC   domain, a catalytic domain, a coiled-coil structure region, a
CC   proline-rich region, a spacer region, or a C-terminal tail. Also
CC   include are the encoding polynucleotides. The novel kinase is useful for
CC   identifying a substance that modulates the activity of a kinase
CC   The polynucleotide is useful for detecting nucleic acid encoding a kinase
CC   polypeptide in a sample. The kinases, polynucleotides modulators
CC   of the kinases and anti-kinase antibodies are useful for diagnosing and
CC   treating a disease or disorder such as cancers, immune-related disease
CC   and disorders, cardiovascular disease, brain or neuronal-associated
CC   diseases, and metabolic disorders, e.g. cancers of tissues, cancers of
CC   haematopoietic origin, diseases of the central or peripheral nervous
CC   system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,
CC   amyotrophic lateral sclerosis, viral infections, infections caused by
CC   prions, infections caused by bacteria, infections caused by fungi, or
CC   ocular diseases, migraine, pain, sexual dysfunction, mood disorders,
CC   attention disorders, cognition disorders, hypotension, hypertension,
CC   psychotic disorders, neurological disorders, dyskinesias, metabolic
CC   disorders, or organ transplant rejection and many other diseases
CC   and disorders given in the specification. The polynucleotide is
CC   useful in gene therapy techniques. The two novel kinases are
CC   designated SGK341 (the gene is located on chromosome Xp22.1) and SGK351
CC   (chromosome 17q22-25). The present sequence is a PCR primer used to
CC   investigate the tissue specific expression of SGK351.
XX
SQ   Sequence 25 BP; 10 A; 2 C; 6 G; 7 T; 0 other;

  Query Match             1.1%;  Score 25;  DB 23;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 6.8e+03;
  Matches   25;  Conservative    0;  Mismatches   0;  Indels    0;  Gaps    0;

Qy    560 GAGAACTATTTATGCAGTTAGAAAG 584
          |||||||||||||||||||||||||
Db      1 GAGAACTATTTATGCAGTTAGAAAG 25


RESULT 10
AAS19269/c
ID   AAS19269 standard; DNA; 25 BP.
XX
AC   AAS19269;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human protein kinases SGK351 PCR primer #2.
XX
KW   Human; ss; SGK351; protein kinase; cancers; immune-related disease;
KW   cardiovascular disease; neuronal-associated disease; metabolic disorder;
KW   tissue cancer; haematopoietic cancer; central nervous system disorder;
KW   peripheral nervous system disorder; Alzheimer's disease;
KW   Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW   viral infection; prion infection; bacterial infection; fungal infection;
KW   ocular disease; migraine; pain; sexual dysfunction; mood disorder;
KW   attention disorder; cognition disorder; hypotension; hypertension;
KW   psychotic disorder; neurological disorder; dyskinesia; gene therapy;
KW   metabolic disorder; organ transplant rejection; PCR primer.
XX
OS   Homo sapiens.
XX
PN   WO200177338-A2.
XX
PD   18-OCT-2001.
XX
PF   10-APR-2001; 2001WO-US11675.
XX
PR   10-APR-2000; 2000US-195953P.
PR   01-MAY-2000; 2000US-201015P.
PR   22-JUN-2000; 2000US-213805P.
XX
PA   (SUGE-) SUGEN INC.
XX
PI   Plowman G,  Whyte D,  Manning G,  Sudarsanam S,  Martinez R;
PI   Caenepeel S;
XX
DR   WPI; 2001-657173/75.
XX
PT   Novel kinase polypeptides useful for identifying modulators of kinase
PT   polypeptide that are used for inhibiting kinase activity for treating
PT   cancers, Alzheimer's disease, Parkinson's disease, multiple sclerosis
PT   -
XX
PS   Example 2b; Page 104; 167pp; English.
XX
CC   The invention relates to an isolated, enriched, or purified kinase
CC   polypeptide comprising a 1360 or 198 residue amino acid sequence, fully
CC   defined in specification, or comprising a kinase except which
CC   lacks one or more of N- or C-terminal domain, C-terminal catalytic
CC   domain, a catalytic domain, a coiled-coil structure region, a
CC   proline-rich region, a spacer region, or a C-terminal tail. Also
```

```
CC   include are the encoding polynucleotides. The novel kinase is useful for
CC   identifying a substance that modulates the activity of a kinase
CC   The polynucleotide is useful for detecting nucleic acid encoding a kinase
CC   polypeptide in a sample. The kinases, polynucleotides modulators
CC   of the kinases and anti-kinase antibodies are useful for diagnosing and
CC   treating a disease or disorder such as cancers, immune-related disease
CC   and disorders, cardiovascular disease, brain or neuronal-associated
CC   diseases, and metabolic disorders, e.g. cancers of tissues, cancers of
CC   haematopoietic origin, diseases of the central or peripheral nervous
CC   system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,
CC   amyotrophic lateral sclerosis, viral infections, infections caused by
CC   prions, infections caused by bacteria, infections caused by fungi, or
CC   ocular diseases, migraine, pain, sexual dysfunction, mood disorders,
CC   attention disorders, cognition disorders, hypotension, hypertension,
CC   psychotic disorders, neurological disorders, dyskinesias, metabolic
CC   disorders, or organ transplant rejection and many other diseases
CC   and disorders given in the specification. The polynucleotide is
CC   useful in gene therapy techniques. The two novel kinases are
CC   designated SGK341 (the gene is located on chromosome Xp22.1) and SGK351
CC   (chromosome 17q22-25). The present sequence is a PCR primer used to
CC   investigate the tissue specific expression of SGK351.
XX
SQ   Sequence 25 BP; 6 A; 6 C; 5 G; 8 T; 0 other;

  Query Match             1.1%;  Score 25;  DB 23;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 6.8e+03;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1087 CACATTAACTGGGAAGAACTTCTGG 1111
           |||||||||||||||||||||||||
Db     25 CACATTAACTGGGAAGAACTTCTGG 1

RESULT 11
ABN59017
ID   ABN59017 standard; DNA; 60 BP.
XX
AC   ABN59017;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:31765.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 31765; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 9 A; 14 C; 14 G; 23 T; 0 other;

  Query Match             1.0%;  Score 24.4;  DB 24;  Length 60;
  Best Local Similarity   68.0%;  Pred. No. 1.5e+04;
  Matches   34;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   1223 CTCTCAGTGAAAGTGCCAATCAGGTCTTTCTGGGTTTTACATATGTGGCT 1272
           | |  ||||  | |||  ||||| | |||| ||| ||||  | ||| |||
Db     11 CACCTAGTGCTAATGCACATCAGCTTTTTCGGGGGTTTAGTTTTGTTGCT 60

RESULT 12
AAV76048
ID   AAV76048 standard; DNA; 57 BP.
XX
AC   AAV76048;
XX
DT   16-MAR-1999  (first entry)
XX
DE   Staphylococcus aureus contig SEQ ID #1737.
XX
KW   Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW   cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW   skin infection; surgical wound infection; scalded skin syndrome;
KW   toxic shock syndrome; ds.
XX
OS   Staphylococcus aureus.
XX
PN   EP786519-A2.
XX
PD   30-JUL-1997.
XX
PF   07-JAN-1997;   97EP-0100117.
XX
PR   05-JAN-1996;   96US-0009861.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Barash SC,  Choi GH,  Dillon PJ,  Fannon MR,  Kunsch CA;
PI   Rosen CA;
XX
DR   WPI; 1997-374922/35.
XX
PT   Polynucleotide(s) and proteins derived from Staphylococcus aureus  -
PT   stored on computer readable medium and used in the production of
PT   anti-S.aureus vaccines
XX
PS   Claim 1; Page 2046; 3271pp; English.
XX
CC   This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC   of the invention. The DNA sequences are recorded on a computer readable
```

```
CC   medium, preferably selected from a floppy or hard disk, random access
CC   memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC   the S.aureus DNA sequences allows putative functions to be assigned so
CC   that protein-encoding or regulatory regions of commercial, therapeutic or
CC   industrial importance can be obtained. Specifically, sequences which are
CC   likely to encode antigens have been identified and these polypeptides can
CC   be used in a vaccine composition against S.aureus infection. The
CC   polypeptides can also be used in a kit for the immunodetection of
CC   S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC   including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC   skin and surgical wound infections, scalded skin syndrome, toxic shock
CC   syndrome, etc. Organisms transformed with the DNA sequences can be used
CC   for recombinant production of the polypeptides. The new DNA sequences
CC   (and their fragments) are useful as primers or probes for isolating
CC   homologues of any of the S.aureus DNA sequences contained on the
CC   computer readable medium.
XX
SQ   Sequence 57 BP; 29 A; 6 C; 10 G; 12 T; 0 other;

  Query Match             1.0%;  Score 24.2;  DB 18;  Length 57;
  Best Local Similarity   71.1%;  Pred. No. 1.6e+04;
  Matches   32;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    421 AGAAATGCTAAAGATACAGCTCATACAAAAGCAGAACGGAATATT 465
          | | | ||||||| | || | ||| || |||| ||  || ||||
Db      1 AAAGAAGCTAAAGAAAAAGTTGATAAAACAGCAAAAGTGATTATT 45


RESULT 13
AAT32212
ID   AAT32212 standard; DNA; 30 BP.
XX
AC   AAT32212;
XX
DT   26-SEP-1996  (first entry)
XX
DE   Rat mitogen-activated S6 kinase p70s6kdeltaN PCR primer.
XX
KW   PCR primer; polymerase chain reaction; amplification; construction;
KW   transcription system; measurement; interaction; phosphatase; kinase
KW   auto-inhibitory domain; identification; antagonist; agonist;
KW   calcineurin; inhibition; T-cell activation; prevention; cancer;
KW   graft rejection; treatment; arthritis; autoimmune disease; allergy;
KW   psoriasis; Alzheimer's disease; regulation; cell proliferation;
KW   rat; mitogen-activated S6 kinase; ss.
XX
OS   Synthetic.
XX
PN   WO9603501-A1.
XX
PD   08-FEB-1996.
XX
PF   12-JUL-1995;   95WO-EP02724.
XX
PR   22-JUL-1994;   94EP-0810435.
XX
PA   (CIBA ) CIBA GEIGY AG.
XX
PI   Chaudhuri B,  Furst P,  Stephan C,  Fuerst P;
XX
DR   WPI; 1996-117046/12.
XX
PT   Transcription system for measuring interaction between phosphatase
PT   or kinase and its auto:inhibitory domain - used to identify agonist
PT   or antagonists of kinase(s) useful e.g. in preventing graft
PT   rejection
XX
PS   Example 16; Page 66; 83pp; English.
XX
CC   The primer pair AAT32210/12 was used for the PCR amplification of
CC   the rat mitogen-activated S6 kinase, p70s6kdeltaD. The PCR prod.
CC   was used in the construction of a transcription system (TS), for
CC   measuring interaction between a phosphatase or kinase (PoK),
CC   including muteins or fragments which bind an autoinhibitory domain
CC   (AID), and an AID. The TS comprises the DNA binding domain (DBD)
CC   of a transcription factor (TF), and separately the transcription
CC   activation domain of a TF, where the domains are respectively
CC   linked to a 1st polypeptide comprising a PoK (or an AID binding
CC   fragment), and a 2nd polypeptide comprising an AID able to bind
CC   the 1st polypeptide. The TS can be used for the identification of
CC   PoK (ant)agonists, e.g. calcineurin (ant)agonists which can be
CC   used inhibit T-cell activation, e.g. to prevent graft rejection.
CC   Other possible applications of the (ant)agonists are in the
CC   treatment of cancer, arthritis, psoriasis, autoimmune disease,
CC   allergy and Alzheimer's disease, and in the regulation of cell
CC   proliferation.
XX
SQ   Sequence 30 BP; 7 A; 5 C; 8 G; 10 T; 0 other;

  Query Match             1.0%;  Score 24;  DB 17;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    289 CCAGAATGTTTTGAGCTACTTCGG 312
          ||||||||||||||||||||||||
Db      7 CCAGAATGTTTTGAGCTACTTCGG 30


RESULT 14
AAT32213
ID   AAT32213 standard; DNA; 32 BP.
XX
AC   AAT32213;
XX
DT   26-SEP-1996  (first entry)
XX
DE   Yeast ACE1-rat mitogen-activated S6 kinase fusion gene PCR primer.
XX
KW   PCR primer; polymerase chain reaction; amplification; construction;
KW   transcription system; measurement; interaction; phosphatase; kinase
KW   auto-inhibitory domain; identification; antagonist; agonist;
KW   calcineurin; inhibition; T-cell activation; prevention; cancer;
KW   graft rejection; treatment; arthritis; autoimmune disease; allergy;
KW   psoriasis; Alzheimer's disease; regulation; cell proliferation;
KW   rat; mitogen-activated S6 kinase; yeast; transcriptional activatot;
KW   fusion; ss.
XX
OS   Synthetic.
XX
PN   WO9603501-A1.
XX
PD   08-FEB-1996.
XX
PF   12-JUL-1995;   95WO-EP02724.
XX
PR   22-JUL-1994;   94EP-0810435.
XX
PA   (CIBA ) CIBA GEIGY AG.
XX
PI   Chaudhuri B,  Furst P,  Stephan C,  Fuerst P;
XX
DR   WPI; 1996-117046/12.
XX
PT   Transcription system for measuring interaction between phosphatase
PT   or kinase and its auto:inhibitory domain - used to identify agonist
PT   or antagonists of kinase(s) useful e.g. in preventing graft
PT   rejection
XX
PS   Example 17; Page 67; 83pp; English.
XX
CC   The primer pair AAT32210/13 was used for the PCR amplification of a
CC   yeast transcriptional activator (ACE1)/rat mitogen-activated S6
CC   kinase, fusion gene. The PCR prod. was used in the construction of
CC   a transcription system (TS), for measuring interaction between a
CC   phosphatase or kinase (PoK), including muteins or fragments which
CC   bind an autoinhibitory domain (AID), and an AID. The TS comprises
```

CC the DNA binding domain (DBD) of a transcription factor (TF), and
CC separately the transcription activation domain of a TF, where the
CC domains are respectively linked to a 1st polypeptide comprising a
CC PoK (or an AID binding fragment), and a 2nd polypeptide comprising
CC an AID able to bind the 1st polypeptide. The TS can be used for
CC the identification of PoK (ant)agonists, e.g. calcineurin
CC (ant)agonists which can be used inhibit T-cell activation, e.g. to
CC prevent graft rejection. Other possible applications of the
CC (ant)agonists are in the treatment of cancer, arthritis, psoriasis,
CC autoimmune disease, allergy and Alzheimer's disease, and in the
CC regulation of cell proliferation.
XX
SQ Sequence 32 BP; 12 A; 4 C; 5 G; 11 T; 0 other;

Query Match 1.0%; Score 23.8; DB 17; Length 32;
Best Local Similarity 92.6%; Pred. No. 1.6e+04;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1290 TGTGAAAGAAAAGTTTTCCTTTGAACC 1316
||||||||||||||||| ||||||||
Db 6 TCTGAAAGAAAAGTTTTCTTTGAACC 32

RESULT 15
AAL29254/c
ID AAL29254 standard; DNA; 49 BP.
XX
AC AAL29254;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #2462.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 2088; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 49 BP; 3 A; 3 C; 4 G; 39 T; 0 other;

Query Match 1.0%; Score 23.8; DB 22; Length 49;
Best Local Similarity 72.1%; Pred. No. 1.9e+04;
Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1760 AAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
||| || |||| ||||| | ||||||||| ||| |||| |
Db 48 AAATAAACAATAAATGGGCCCAAAAAAAAAAAAAAAAAAAAAA 6

RESULT 16
ABN45373/c
ID ABN45373 standard; DNA; 60 BP.
XX
AC ABN45373;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:18121.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 18121; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological

CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 8 A; 10 C; 13 G; 29 T; 0 other;

Query Match 1.0%; Score 23.8; DB 24; Length 60;
Best Local Similarity 62.7%; Pred. No. 2.1e+04;
Matches 37; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 246 AATCTCAGAAACTAGTGTGAACAGAGGGCCAGAAAAAATCAGACCAGAATGTTTTGAGC 304
|| |||| |||||| |||| ||| ||| |||||| | || ||| |||
Db 59 AACAGCAGACACTAGTAAAAACAAAGGCACAGTAAAAATTGAGACCCAAAATTTGCAGC 1

RESULT 17
ABN50683
ID ABN50683 standard; DNA; 60 BP.
XX
AC ABN50683;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:23431.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 23431; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological

CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 19 A; 11 C; 17 G; 13 T; 0 other;

Query Match 1.0%; Score 23.8; DB 24; Length 60;
Best Local Similarity 66.7%; Pred. No. 2.1e+04;
Matches 34; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 2180 TTTAAAATGCGAAATTATTGGTTGGTGTGAAGAAAGCCAGACAACTTCTGT 2230
|||| | || | || ||||| | ||| | |||| ||| |||| || |
Db 9 TTTACACTGTGCAAGTATTGAGAAGAGTGCATAAAGACAGGGAACTACTCT 59

RESULT 18
ABL00050
ID ABL00050 standard; DNA; 50 BP.
XX
AC ABL00050;
XX
DT 05-MAR-2002 (first entry)
XX
DE Human silent noncoding SNP oligonucleotide SEQ ID NO:41.
XX
KW Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;
KW immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;
KW autoimmune disease; inflammation; cancer; nervous system disease;
KW infection; polymorphic protein; ds.
XX
OS Homo sapiens.
XX
PN WO200138586-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32311.
XX
PR 24-NOV-1999; 99US-0167383.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-355949/37.
XX
PT Isolated human nucleic acids comprising one or more single nucleotide
PT polymorphisms, useful for treating a subject suffering from a
PT pathology, e.g. autoimmune diseases, ascribed to the presence of a
PT sequence polymorphism -
XX
PS Claim 1; Page 257; 674pp; English.
XX
CC ABL00010 to ABL01104 represent human nucleic acid oligonucleotides
CC comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
CC to ABB56903 represent human peptides encoded by some of the SNP
CC oligonucleotides. The sequences from the present invention can have
CC immunosuppressive, cytostatic, antiinflammatory, neuroprotective and
CC antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC and antibodies from the present invention can be used for treating a
CC subject suffering from, at risk for, or suspected of, suffering from a
CC pathology ascribed to the presence of a sequence polymorphism. The
CC pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC the nervous system, and infection by pathogenic microorganisms. The SNPs
CC are also useful for determining which forms of a characterised
CC polymorphism are present in individuals. The antibodies may be used in
CC the detection, quantitation and/or cellular or tissue localisation of a
CC polymorphic protein (e.g., for use in measuring levels of the

```
CC   polymorphic protein within appropriate physiological samples).
XX
SQ   Sequence 50 BP; 22 A; 5 C; 3 G; 20 T; 0 other;

  Query Match             1.0%;  Score 23.6;  DB 23;  Length 50;
  Best Local Similarity   76.3%;  Pred. No. 2.2e+04;
  Matches   29;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy   2117 TATATAAATATATATTTTTCAAATAGATTTTTGATTCA 2154
           |||||| |||||| || | |||| | || ||| || ||
Db     13 TATATATATATATTTTATCCAAAAATATGTTTTATACA 50


RESULT 19
AAL28430/c
ID   AAL28430 standard; DNA; 51 BP.
XX
AC   AAL28430;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #1638.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 1849; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 8 A; 2 C; 1 G; 40 T; 0 other;

  Query Match             1.0%;  Score 23.6;  DB 22;  Length 51;
  Best Local Similarity   69.6%;  Pred. No. 2.2e+04;
  Matches   32;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1757 TTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
           ||| |||||| || || || |  ||| |||||   |||  |||| |
Db     51 TTTTAAAAATAAAACATTGATAAAAAGAAAAAAAAAAAAAAAAAA 6


RESULT 20
AAH84250
ID   AAH84250 standard; cDNA; 60 BP.
XX
AC   AAH84250;
XX
DT   21-SEP-2001  (first entry)
XX
DE   Human cell death protective cDNA clone CNI-00717 ORF24, SEQ:231.
XX
KW   Cell death protective; apoptosis; necrosis; human; drug screening;
KW   cell death-associated disorder; central nervous system disorder;
KW   psychiatric disorder; neurological disorder; ischaemia-related disorder;
KW   stroke; cerebral infarction; ischaemic encephalopathy;
KW   neurodegenerative disorder; Alzheimer's disease; Huntington's disease;
KW   Parkinson's disease; infection; meningitis; malaria; trypanosomiasis;
KW   vascular disease; opthalmological disorder; diabetic retinopathy;
KW   macular degeneration; hypertension; myocardial infarction;
KW   atherosclerosis; respiratory disorder; asthma; transgenic animal;
KW   chronic obstructive pulmonary disease; neoplastic condition; cancer;
KW   benign tumour; anaemia; gastrointestinal disorder; gastritis;
KW   ulcerative colitis; liver disease; biliary cirrhosis; kidney disorder;
KW   glomerulonephritis; cystitis; endometriosis; endocrine disorder;
KW   Grave's disease; Hashimoto's thyroiditis; skin condition; dermatitis;
KW   urticaria; immune disorder; acquired immunodeficiency syndrome; AIDS;
KW   open reading frame; ORF; ss.
XX
OS   Homo sapiens.
XX
PN   WO200145638-A2.
XX
PD   28-JUN-2001.
XX
PF   11-DEC-2000; 2000WO-US33547.
XX
PR   14-DEC-1999;   99US-0461697.
XX
PA   (COGE-) COGENT NEUROSCIENCE INC.
XX
PI   Lo DC,  Barney S,  Thomas MB,  Portbury SD,  Puranam K,  Katz LC;
XX
DR   WPI; 2001-390297/41.
DR   P-PSDB; AAG98720.
XX
PT   Novel protective sequence polynucleotides and polypeptides, used to
PT   identify modulators of their expression and activity, which are used in
PT   to treat central nervous system conditions, diseases and disorders -
XX
PS   Claim 2; Fig 9X; 325pp; English.
XX
CC   Sequences AAH84132-AAH84370 represent human nucleic acid sequences which
CC   protect against cell death (i.e., apoptosis or necrosis). Sequences
CC   AAH84132, AAH84145, AAH84170, AAH84201, AAH84210, AAH84226, AAH84265,
CC   AAH84281, AAH84315 and AAH84367 represent 10 full-length cDNA clones,
CC   while the remaining nucleic acid sequences within the range given above
CC   represent the open reading frames (ORFs) of these cDNA clones. Sequences
CC   AAG98610-AAG98829 represent the polypeptides encoded by the cell death
CC   protective ORFs. The cell death protective cDNA clones are able to
CC   prevent, delay or reverse progression through the apoptotic or necrotic
CC   pathways when injected into a cell predisposed to or undergoing cell
CC   death. The cell death protective nucleic acids and polypeptides can be
CC   used in the diagnosis and treatment of disorders associated with cell
CC   death, and to screen for compounds which modulate their activity or
```

```
CC   expression. Such modulators, preferably a small organic molecule, an
CC   antibody, a ribozyme, or an antisense molecule, can also be used to treat
CC   cell death-related diseases. Such diseases include those associated with
CC   the central nervous system including psychiatric or neurological
CC   disorders, especially ischaemia-related conditions such as strokes, and
CC   also includes neurodegenerative disorders such as Alzheimer's disease,
CC   Huntington's disease, or Parkinson's disease. The modulators may also be
CC   used to treat infections such as meningitis, malaria, or trypanosomiasis;
CC   vascular diseases such as ischaemic encephalopathy or cerebral
CC   infarction; eye conditions such as diabetic retinopathy or macular
CC   degeneration; hypertension; myocardial infarction; atherosclerosis;
CC   respiratory conditions such as asthma or chronic obstructive pulmonary
CC   disease; neoplastic conditions such as cancers or benign tumours; blood
CC   cell conditions such as anaemia; gastrointestinal conditions such as
CC   gastritis or ulcerative colitis; liver conditions such as biliary
CC   cirrhosis; kidney disorders such as glomerulonephritis; cystitis;
CC   endometriosis; endocrine disorders such as Grave's disease or Hashimoto's
CC   thyroiditis; skin conditions such as dermatitis or urticaria; or immune
CC   system disorders such as acquired immunodeficiency syndrome (AIDS). The
CC   nucleic acids may additionally be used to generate animal models of
CC   cell death-associated disorders. The present sequence represents a
CC   cell death protective ORF.
XX
SQ   Sequence 60 BP; 31 A; 1 C; 20 G; 8 T; 0 other;

  Query Match             1.0%;  Score 23.6;  DB 22;  Length 60;
  Best Local Similarity   64.8%;  Pred. No. 2.4e+04;
  Matches   35;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy   1635 AAAAAGGTGGAGAGGGAGATGTGTGAGCATCCTGCAAGGTGAAACAAGACTCAA 1688
           || ||| || ||  ||| ||| |  || | |  | || | |||| |||| | ||
Db      7 AAGAAGATGAAGGTGGAAATGAGGAAGAAGCTGGAAAAGAGAAAGAAGATTTAA 60


RESULT 21
AAS19340
ID   AAS19340 standard; DNA; 58 BP.
XX
AC   AAS19340;
XX
DT   20-MAR-2002  (first entry)
XX
DE   Oligonucleotide 5981 used to contruct plasmid XL2725.
XX
KW   ss; DNA purification; triple helix; plasmid purification;
KW   oligonucleotide 5981; XL2725.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   repeat_region   6..56
FT                   /*tag=  a
FT                   /rpt_type= "TANDEM"
FT   repeat_unit     6..8
FT                   /*tag=  b
FT                   /note= "GGA repeat type"
XX
PN   WO200192511-A2.
XX
PD   06-DEC-2001.
XX
PF   25-MAY-2001; 2001WO-US17122.
XX
PR   26-MAY-2000; 2000US-0580923.
XX
PA   (AVET ) AVENTIS PHARMA SA.
XX
PI   Crouzet J,  Scherman D,  Wils P,  Blanche F,  Cameron B;
XX
DR   WPI; 2002-097772/13.
XX
PT   Purifying double-stranded (ds) DNA from a solution containing dsDNA and
PT   other components, comprises passing the solution through a support
PT   comprising a covalently coupled oligonucleotide able to form a triple
PT   helix with the dsDNA -
XX
PS   Example 7.1; Page 20; 40pp; English.
XX
CC   This invention comprises a method of purifying double-stranded DNA from
CC   a solution containing the double-stranded DNA mixed with other
CC   components, comprising passing the solution through a support comprising
CC   a covalently coupled oligonucleotide capable of forming a triple helix
CC   with the double-stranded DNA by hybridisation with a specific sequence
CC   present in the double-stranded DNA. The method is useful for purifying
CC   double-stranded DNA contained in a solution and mixed with other
CC   components. The new method is a simple, rapid and effective method for
CC   DNA purification, and makes it possible to obtain especially high
CC   purities with high yields. The method enables DNA to be purified from
CC   complex mixtures comprising other nucleic acids, proteins, endotoxins,
CC   nucleases and the like. The supports may be readily recycled, and the
CC   DNAs obtained display improved properties to pharmaceutical safety.
CC   Further, the method entails only one step contrary to prior art.
CC   The present sequence represents an oligonucleotide 5981 used to
CC   create the XL2725 plasmid which was used in an example of the DNA
CC   purification method of the invention.
XX
SQ   Sequence 58 BP; 18 A; 2 C; 37 G; 1 T; 0 other;

  Query Match             1.0%;  Score 23.4;  DB 24;  Length 58;
  Best Local Similarity   63.2%;  Pred. No. 2.7e+04;
  Matches   36;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy    118 GACCTGGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAG 174
          || | |||  ||   |||| | ||     ||||| |||  |||||||| ||   ||
Db      1 GATCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 57


RESULT 22
AAT22872/c
ID   AAT22872 standard; cDNA to mRNA; 50 BP.
XX
AC   AAT22872;
XX
DT   29-AUG-1996  (first entry)
XX
DE   Human gene signature HUMGS04557.
XX
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
XX
OS   Homo sapiens.
XX
PN   WO9514772-A1.
XX
PD   01-JUN-1995.
XX
PF   11-NOV-1994;   94WO-JP01916.
XX
PR   12-NOV-1993;   93JP-0355504.
XX
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
XX
PI   Matsubara K,  Okubo K;
XX
DR   WPI; 1995-206931/27.
XX
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
XX
PS   Claim 1; Page 1231; 2245pp; Japanese.
XX
CC   A single-stranded DNA (or its complementary strand or the corresp.
```

CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in AAT19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
XX
SQ Sequence 50 BP; 25 A; 4 C; 4 G; 17 T; 0 other;

Query Match 1.0%; Score 23.2; DB 16; Length 50;
Best Local Similarity 77.8%; Pred. No. 2.8e+04;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2104 TGAATCTTTTTTTTATATAAATATATATTTTTCAAA 2139
Db 41 TCAGTATTTTTAATACATAAATATATATTTTACATA 6

RESULT 23
AAL28930/c
ID AAL28930 standard; DNA; 50 BP.
XX
AC AAL28930;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #2138.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 1994; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 50 BP; 8 A; 5 C; 4 G; 33 T; 0 other;

Query Match 1.0%; Score 23.2; DB 22; Length 50;
Best Local Similarity 70.5%; Pred. No. 2.8e+04;
Matches 31; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1760 AAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATAG 1803
Db 47 AAAACAACAAAATTTGGTGCTTAAAAAAAGTAAAATAAAAAAAG 4

RESULT 24
AAL31877/c
ID AAL31877 standard; DNA; 51 BP.
XX
AC AAL31877;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #5085.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 2849; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of

```
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 5 A; 4 C; 4 G; 38 T; 0 other;

  Query Match             1.0%;  Score 23.2;  DB 22;  Length 51;
  Best Local Similarity   70.5%;  Pred. No. 2.8e+04;
  Matches   31;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   1759 TAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
           ||||||||||  | ||  |   |||||||||   |||  |||| |
Db     45 TAAAAAATCAGTGAAATGGAAAAAAAAAAAAAAAAAAAAAAAAA 2


RESULT 25
ABN34677
ID   ABN34677 standard; DNA; 60 BP.
XX
AC   ABN34677;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:7425.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 7425; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 17 A; 19 C; 11 G; 13 T; 0 other;

  Query Match             1.0%;  Score 23.2;  DB 24;  Length 60;
  Best Local Similarity   65.4%;  Pred. No. 3.1e+04;
  Matches   34;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy   1538 GGCCATACAAAAAACAAGCTTTTCCCATGATCTCCAAACGGCCAGAGCACCT 1589
           || | | |||||  | ||| || |||  | | |||  ||   |||||| |||
Db      5 GGACGTGCAAAAGTCTAGCCTTGCCCGAGCTTTCCTTACAATCAGAGCCCCT 56


RESULT 26
AAH44691/c
ID   AAH44691 standard; DNA; 41 BP.
XX
AC   AAH44691;
XX
DT   07-DEC-2001  (first entry)
XX
DE   Human type-I aminoacyl tRNA synthetase 10 probe 2 SEQ ID NO:9.
XX
KW   Human; type-I aminoacyl tRNA synthetase 10; malignant tumour;
KW   haemopathy; human immunodeficiency virus; HIV infection;
KW   immunological disease; inflammation; probe; ss.
XX
OS   Homo sapiens.
XX
PN   CN1301715-A.
XX
PD   04-JUL-2001.
XX
PF   27-DEC-1999;   99CN-0125371.
XX
PR   27-DEC-1999;   99CN-0125371.
XX
PA   (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2001-550469/62.
XX
PT   New polypeptide-I-type aminoacyl/tRNA synthetase 10 and encoding
PT   polynucleotide useful for treating tumor, hemopathy, infection and
PT   immunological disease -
XX
PS   Example 7; Page 20 (Disclosure); 32pp; Chinese.
XX
CC   The present invention describes the human type-I aminoacyl tRNA
CC   synthetase 10 protein. Also described are polynucleotides encoding the
CC   type-I amino acyl tRNA synthetase 10 protein, and a DNA recombination
CC   process to produce the protein. The protein can be used for treating
CC   various diseases, such as malignant tumour, haemopathy, human
CC   immunodeficiency virus infection, immunological diseases and various
CC   inflammations. The present sequence represents a probe for type-I
CC   aminoacyl tRNA synthetase 10, which is used in an example from the
CC   present invention.
XX
SQ   Sequence 41 BP; 3 A; 5 C; 4 G; 29 T; 0 other;

  Query Match             1.0%;  Score 22.8;  DB 22;  Length 41;
  Best Local Similarity   79.4%;  Pred. No. 3.2e+04;
  Matches   27;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy   2045 AAAATGGGAAGGCAAAGACAAAAGAAACTTACCA 2078
```

```
          ||||  | || ||||||| ||||| ||||||| |
Db     37 AAAAAAGAAAAGCAAAGAAAAAAGGAACTTACAA 4


RESULT 27
AAZ67514/c
ID   AAZ67514 standard; DNA; 47 BP.
XX
AC   AAZ67514;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human map-related biallelic marker SEQ ID NO:1861.
XX
KW   Human genome; biallelic marker; high density disequilibrium map;
KW   genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW   haplotyping; hybridisation; identification; characterisation;
KW   diagnosis; single nucleotide polymorphism; SNP; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(24,T)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO9954500-A2.
XX
PD   28-OCT-1999.
XX
PF   21-APR-1999;   99WO-IB00822.
XX
PR   21-APR-1998;   98US-0082614.
PR   23-NOV-1998;   98US-0109732.
XX
PA   (GEST ) GENSET.
XX
PI   Cohen D,  Blumenfeld M,  Chumakov I;
XX
DR   WPI; 2000-013267/01.
XX
PT   Novel biallelic markers used to construct a high density disequilibrium
PT   map of the human genome  -
XX
PS   Claim 1; Page 627; 2745pp; English.
XX
CC   AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC   invention, which contain a polymorphic base at position 24 of their
CC   nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC   primers for the biallelic markers. The biallelic markers of the
CC   invention have a variety of uses: they can be used for high density
CC   mapping of the human genome, and in complex association studies and
CC   haplotyping studies which are useful in determining the genetic basis
CC   for disease states. Compositions and methods of the invention can also
CC   be useful for the identification of the targets for the development of
CC   pharmaceutical agents and diagnostic methods, as well as the
CC   characterisation of the differential efficacious responses to and side
CC   effects from pharmaceutical agents acting on a disease as well as other
CC   treatment.
CC   N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297
CC   and 3367, are not actually given a sequence in the Sequence Listing
CC   from the present invention.
XX
SQ   Sequence 47 BP; 17 A; 7 C; 6 G; 17 T; 0 other;

  Query Match             1.0%;  Score 22.8;  DB 21;  Length 47;
  Best Local Similarity   79.4%;  Pred. No. 3.5e+04;
  Matches   27;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy   2217 CAGACAACTTCTGTTTCTTCTCTTGGTGAAATAA 2250
          |||| |||||| |||||||| |||   ||||||
Db     47 CAGAAAACTTCAGTTTCTTCATTTGTAAAAATAA 14


RESULT 28
AAZ23542
ID   AAZ23542 standard; DNA; 48 BP.
XX
AC   AAZ23542;
XX
DT   21-DEC-1999  (first entry)
XX
DE   Human DNA fragment 1.
XX
KW   Assay; amplification; hybridisation; probe; detection; viral; bacterial;
KW   cellular; yeast; fungal; primer; ss.
XX
OS   Homo sapiens.
XX
PN   DE19814828-A1.
XX
PD   07-OCT-1999.
XX
PF   02-APR-1998;   98DE-1014828.
XX
PR   02-APR-1998;   98DE-1014828.
XX
PA   (HOFF ) ROCHE DIAGNOSTICS GMBH.
XX
PI   Kessler C,  Haberhausen G,  Batz H,  Oerum H;
XX
DR   WPI; 1999-552286/47.
XX
PT   Nucleic acid amplification assay for detecting viral, bacterial,
PT   cellular, yeast or fungal nucleic acids  -
XX
PS   Disclosure; Fig 4; 28pp; German.
XX
CC   This invention describes a novel assay for a nucleic acid comprises:
CC   (a) generating amplification products from a fragment of the nucleic
CC   acid, (b) contacting the amplification products with a probe; and
CC   (c) detecting hybridization between the amplification product and the
CC   probe. The assay is useful for detection of viral, bacterial, cellular,
CC   yeast or fungal nucleic acids in human, animal, bacterial, plant, yeast
CC   or fungal samples, e.g. feces, smears, cell suspensions, cultures or
CC   tissue, cell or liquid biopsy samples. This sequence represents a
CC   fragment of the human genome which is used in the method of the
CC   invention.
XX
SQ   Sequence 48 BP; 9 A; 17 C; 14 G; 8 T; 0 other;

  Query Match             1.0%;  Score 22.8;  DB 20;  Length 48;
  Best Local Similarity   71.4%;  Pred. No. 3.5e+04;
  Matches   30;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    874 ATGTATGACATGCTGACTGGAGCACCCCCATTCACTGGGGAG 915
          |||| || | ||| | ||  || ||||||| || | || |||
Db      4 ATGTGTGTCGTGCAGCCTCCAGGACCCCCACTCCCGGGAGAG 45


RESULT 29
ABN71559/c
ID   ABN71559 standard; DNA; 48 BP.
XX
AC   ABN71559;
XX
DT   01-JUL-2002  (first entry)
XX
DE   Streptococcus agalactiae PCR primer SEQ ID NO 10999.
XX
KW   Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW   group A streptococcus; Streptococcus pyogenes; antibacterial; PCR; ss;
KW   antiinflammatory; infection; vaccine; meningitis; gene therapy; primer.
XX
OS   Streptococcus agalactiae.
OS   Synthetic.
```

XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB04789.
XX
PR 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
PI Tettelin H;
XX
DR WPI; 2002-352536/38.
XX
PT New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
PS Claim 8; Page 4489; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins. The present sequence is that of a PCR primer,
CC useful to the invention.
XX
SQ Sequence 48 BP; 15 A; 5 C; 13 G; 15 T; 0 other;

Query Match 1.0%; Score 22.8; DB 24; Length 48;
Best Local Similarity 71.4%; Pred. No. 3.5e+04;
Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1999 ACTAAGGACAAATTAGCATGCAAGCTTGGTCAAACTTTTCCC 2040
| ||| |||||| ||| ||| ||| ||||||| ||||
Db 43 AATAATGACAAAACAGCCTGCTTTTTTGTACAAACTTGTCCC 2

RESULT 30
AAL28956
ID AAL28956 standard; DNA; 50 BP.
XX
AC AAL28956;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #2164.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 2002; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 50 BP; 27 A; 6 C; 3 G; 14 T; 0 other;

Query Match 1.0%; Score 22.8; DB 22; Length 50;
Best Local Similarity 66.0%; Pred. No. 3.6e+04;
Matches 33; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1753 GGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
|||||| ||| ||||| |||||||||| ||| ||||| |
Db 1 GGATTTATTTTAATTTTTCAATCTGAAAAAAAAAAACCCAAAACAAAAAA 50

RESULT 31
AAL32077
ID AAL32077 standard; DNA; 51 BP.
XX
AC AAL32077;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #5285.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX

PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 2908; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 51 BP; 17 A; 5 C; 5 G; 24 T; 0 other;

Query Match 1.0%; Score 22.8; DB 22; Length 51;
Best Local Similarity 71.4%; Pred. No. 3.6e+04;
Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Qy   2120 ATAAATATATATTTTTCAAATAGATTTTTGATTCAGCTCATT 2161
          | ||  || | ||||||| ||||||||||||  |   | |||
Db      7 AGAAGAATTTTTTTTTCAGATAGATTTTTGAAACTTTTTATT 48
```

RESULT 32
AAC78364/c
ID AAC78364 standard; cDNA; 60 BP.
XX
AC AAC78364;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human cancer associated gene sequence SEQ ID NO:758.
XX
KW Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening; ss.
XX
OS Homo sapiens.
XX
PN WO200055350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05882.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587533/55.
DR P-PSDB; AAB44155.
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer -
XX
PS Claim 1; Page 1305; 2352pp; English.
XX
CC AAC77607 to AAC78448 encode the human cancer associated proteins given
CC in AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnerary; immunomodulator;
CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention.
XX
SQ Sequence 60 BP; 10 A; 2 C; 5 G; 38 T; 5 other;

Query Match 1.0%; Score 22.8; DB 21; Length 60;
Best Local Similarity 63.8%; Pred. No. 3.9e+04;
Matches 30; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```
Qy   1756 TTTTAAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
          ||||||||||  ||   |      |||||||||   |||  |||| |
Db     56 TTTTAAAAAAANNAAAANANANAAAAAAAAAAAAAAAAAAAAAAAAA 10
```

RESULT 33
AAZ66546/c
ID AAZ66546 standard; DNA; 47 BP.
XX
AC AAZ66546;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human map-related biallelic marker SEQ ID NO:893.
XX
KW Human genome; biallelic marker; high density disequilibrium map;
KW genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW haplotyping; hybridisation; identification; characterisation;
KW diagnosis; single nucleotide polymorphism; SNP; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(24,G)
FT /*tag= a
FT /standard_name= "single nucleotide polymorphism"

```
XX
PN   WO9954500-A2.
XX
PD   28-OCT-1999.
XX
PF   21-APR-1999;   99WO-IB00822.
XX
PR   21-APR-1998;   98US-0082614.
PR   23-NOV-1998;   98US-0109732.
XX
PA   (GEST ) GENSET.
XX
PI   Cohen D,  Blumenfeld M,  Chumakov I;
XX
DR   WPI; 2000-013267/01.
XX
PT   Novel biallelic markers used to construct a high density disequilibrium
PT   map of the human genome  -
XX
PS   Claim 1; Page 421; 2745pp; English.
XX
CC   AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC   invention, which contain a polymorphic base at position 24 of their
CC   nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC   primers for the biallelic markers. The biallelic markers of the
CC   invention have a variety of uses: they can be used for high density
CC   mapping of the human genome, and in complex association studies and
CC   haplotyping studies which are useful in determining the genetic basis
CC   for disease states. Compositions and methods of the invention can also
CC   be useful for the identification of the targets for the development of
CC   pharmaceutical agents and diagnostic methods, as well as the
CC   characterisation of the differential efficacious responses to and side
CC   effects from pharmaceutical agents acting on a disease as well as other
CC   treatment.
CC   N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297
CC   and 3367, are not actually given a sequence in the Sequence Listing
CC   from the present invention.
XX
SQ   Sequence 47 BP; 19 A; 5 C; 4 G; 19 T; 0 other;

  Query Match             1.0%;  Score 22.6;  DB 21;  Length 47;
  Best Local Similarity   68.9%;  Pred. No. 3.9e+04;
  Matches   31;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    2106 AATCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGA 2150
           | ||||| |  || |   ||| |||||||  |||| || ||| ||
Db      47 ATTCTTTGTAATAAACTCATAAATATTTTAAAAATTGAGTTTGGA 3


RESULT 34
AAZ69126/c
ID   AAZ69126 standard; DNA; 47 BP.
XX
AC   AAZ69126;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human map-related biallelic marker SEQ ID NO:3482.
XX
KW   Human genome; biallelic marker; high density disequilibrium map;
KW   genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW   haplotyping; hybridisation; identification; characterisation;
KW   diagnosis; single nucleotide polymorphism; SNP; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(24,A)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO9954500-A2.
XX
PD   28-OCT-1999.
XX
PF   21-APR-1999;   99WO-IB00822.
XX
PR   21-APR-1998;   98US-0082614.
PR   23-NOV-1998;   98US-0109732.
XX
PA   (GEST ) GENSET.
XX
PI   Cohen D,  Blumenfeld M,  Chumakov I;
XX
DR   WPI; 2000-013267/01.
XX
PT   Novel biallelic markers used to construct a high density disequilibrium
PT   map of the human genome  -
XX
PS   Claim 3; Page 973; 2745pp; English.
XX
CC   AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC   invention, which contain a polymorphic base at position 24 of their
CC   nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC   primers for the biallelic markers. The biallelic markers of the
CC   invention have a variety of uses: they can be used for high density
CC   mapping of the human genome, and in complex association studies and
CC   haplotyping studies which are useful in determining the genetic basis
CC   for disease states. Compositions and methods of the invention can also
CC   be useful for the identification of the targets for the development of
CC   pharmaceutical agents and diagnostic methods, as well as the
CC   characterisation of the differential efficacious responses to and side
CC   effects from pharmaceutical agents acting on a disease as well as other
CC   treatment.
CC   N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297
CC   and 3367, are not actually given a sequence in the Sequence Listing
CC   from the present invention.
XX
SQ   Sequence 47 BP; 21 A; 2 C; 1 G; 23 T; 0 other;

  Query Match             1.0%;  Score 22.6;  DB 21;  Length 47;
  Best Local Similarity   68.9%;  Pred. No. 3.9e+04;
  Matches   31;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    2100 AACCTGAATCTTTTTTTTATATAAATATATATTTTTCAAATAGAT 2144
           ||  | ||| || |  ||| | ||||||||||||||  | |  ||
Db      47 AATATAAATATTGTAGTTAAAAAAATATATATTTTTATATTTAAT 3


RESULT 35
AAH38300/c
ID   AAH38300 standard; DNA; 51 BP.
XX
AC   AAH38300;
XX
DT   14-AUG-2001  (first entry)
XX
DE   Human SNP flanking oligonucleotide SEQ ID 1096.
XX
KW   Single nucleotide polymorphism; SNP; single nucleotide primer extension;
KW   SNPE; genotyping; agammaglobulinaemia; diabetes insipidus; cancer;
KW   Lesch-Nyhan syndrome; muscular dystrophy; familial hypercholesterolaemia;
KW   polycystic kidney disease; osteogenesis imperfecta; autoimmune disease;
KW   acute intermittent porphyria; rheumatoid arthritis; multiple sclerosis;
KW   inflammation; forensic investigation; paternity analysis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200129262-A2.
XX
PD   26-APR-2001.
XX
PF   13-OCT-2000; 2000WO-US28436.
XX
PR   15-OCT-1999;   99US-0160096.
XX
```

PA (ORCH-) ORCHID BIOSCIENCES INC.
XX
PI Picoult-Newburg L, Pohl M;
XX
DR WPI; 2001-290930/30.
XX
PT New genotyping oligonucleotide, useful for detecting the presence,
PT absence or identity of single polynucleotide polymorphism in a nucleic
PT acid sample -
XX
PS Claim 1; Page 55; 83pp; English.
XX
CC Sequences AAH37205 - AAH40944 represent PCR primers, single nucleotide
CC primer extension (SNPE) primers, and the sequences of regions flanking
CC sites of single nucleotide polymorphisms SNPs. The present invention
CC includes kits for determining the presence or absence of a SNP, using the
CC oligonucleotides of the invention. The PCR primers are used to amplify a
CC SNP flanking sequence, the SNPE primer is used as a genotyping primer.
CC The oligonucleotides are useful for genotyping a nucleic acid sample by
CC performing a single-nucleotide primer extension reaction. The
CC oligonucleotides are useful for determining the presence, absence or
CC identity of a SNP and for genotyping nucleic acid samples, for e.g. to
CC assess by association analysis the genotype of an individual or group of
CC individuals, having a pathological phenotypic trait suspected of being
CC caused by one or more SNPs. Phenotypic traits include diseases e.g.
CC agammaglobulinaemia, diabetes insipidus, Lesch-Nyhan syndrome, muscular
CC dystrophy, familial hypercholesterolaemia, polycystic kidney disease,
CC osteogenesis imperfecta and acute intermittent porphyria. Phenotypic
CC traits also include symptoms of or susceptibility to multifactorial
CC disease of which a component is or may be genetic such as autoimmune
CC diseases, including, rheumatoid arthritis, multiple sclerosis,
CC inflammation, cancer, nervous system diseases and infection by pathogenic
CC microorganism. The method is also useful in forensic investigations and
CC paternity analysis. The present sequence represents a fragment of human
CC DNA flanking the site of a single nucleotide polymorphism.
XX
SQ Sequence 51 BP; 13 A; 9 C; 8 G; 20 T; 1 other;

Query Match 1.0%; Score 22.6; DB 22; Length 51;
Best Local Similarity 68.9%; Pred. No. 4e+04;
Matches 31; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 406 AAGGCAATGATAGTAAGAAATGCTAAAGATACAGCTCATACAAAA 450
|| |||||||||| | ||| | || ||||| | || | ||||
Db 51 ATGGCAATGATATTCAGACACATTACAGATATATCTGCTGGAAAA 7

RESULT 36
AAV65911
ID AAV65911 standard; DNA; 54 BP.
XX
AC AAV65911;
XX
DT 02-FEB-1999 (first entry)
XX
DE HCMV target DNA for invasive cleavage.
XX
KW Nucleic acid detection; multiple sequential invasive cleavage;
KW Invader-directed cleavage assay; nuclease; FEN-1; HCMV; ss.
XX
OS Human cytomegalovirus.
XX
PN WO9842873-A1.
XX
PD 01-OCT-1998.
XX
PF 24-MAR-1998; 98WO-US05809.
XX
PR 24-MAR-1997; 97US-0823516.
XX
PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
PI Brow MAD, Hall JG, Kwiatkowski RW, Lyamichev VI;
PI Mast AL, Vavra SH;
XX
DR WPI; 1998-557036/47.
XX
PT Detecting target nucleic acid by sequence-specific cleavage of
PT complex with two specific oligonucleotides - used to detect
PT cytomegalovirus DNA
XX
PS Example 48; Page 358; 524pp; English.
XX
CC This nucleotide sequence corresponds to nucleotides 3057-3110 of
CC human cytomegavirus (HCMV) genomic DNA. It was used as a target
CC to demonstrate invasive cleavage using Invader oligonucleotide
CC 89-44 (AAV65909), fluorescein-labelled probe 89-76 (see AAV65910) and
CC Pfu FEN-1 nuclease (see AAW79970). The Invader-directed cleavage
CC reaction and the sequential invader-directed cleavage reaction of
CC the present invention provide direct detection methods that combine
CC the advantages of direct detection assays with the specificity
CC provided by a dual or tri oligonucleotide hybridisation assay. The
CC invention relates to means for the detection and characterisation
CC of nucleic acid sequences, and variations in nucleic acid sequences.
XX
SQ Sequence 54 BP; 14 A; 7 C; 25 G; 8 T; 0 other;

Query Match 1.0%; Score 22.6; DB 19; Length 54;
Best Local Similarity 75.7%; Pred. No. 4.2e+04;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 128 AGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGA 164
|| |||||| | ||| || |||| |||| ||||||
Db 4 AGGAAGAGGAGGAGGGTGCTCAGGAGGAGCGGGAGGA 40

RESULT 37
ABN35460
ID ABN35460 standard; DNA; 60 BP.
XX
AC ABN35460;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:8208.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 8208; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple

CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 15 A; 11 C; 18 G; 16 T; 0 other;

Query Match 1.0%; Score 22.6; DB 24; Length 60;
Best Local Similarity 64.2%; Pred. No. 4.4e+04;
Matches 34; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1030 CGTCTGGGAGCTGGTCCTGGGGACGCTGGAGAAGTTCAAGCTCATCCATTCTT 1082
|||||||| ||| |||| | || | || | ||| ||| |||| ||
Db 2 CGTCTGGGTTGTGGACCTGAAGGCGAACGTGATATCAAAGAGCATGCATTTTT 54

RESULT 38
ABN46663/c
ID ABN46663 standard; DNA; 60 BP.
XX
AC ABN46663;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:19411.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 19411; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 10 A; 18 C; 14 G; 18 T; 0 other;

Query Match 1.0%; Score 22.6; DB 24; Length 60;
Best Local Similarity 64.2%; Pred. No. 4.4e+04;
Matches 34; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1634 CAAAAAGGTGGAGAGGGAGATGTGTGAGCATCCTGCAAGGTGAAACAAGACTC 1686
|| ||||||||||| ||| || | || ||| | |||||||| |
Db 60 CATAAAGGTGGAGTCCCTCATGAATGGGATTCGTGCCCGAGTAAACAAGACCC 8

RESULT 39
AAQ69313
ID AAQ69313 standard; DNA; 44 BP.
XX
AC AAQ69313;
XX
DT 22-FEB-1995 (first entry)
XX
DE Human sodium/potassium ATPase alpha 3 subunit gene, target region.
XX
KW DNA protein-binding assay; test sequence; screening sequence;
KW promoter; target; TATA box; Herpes Simplex Virus; HSV;
KW origin of replication; UL9; transcription factor; TFIID:
KW ATP1 A3; ds.
XX
OS Synthetic.
XX
PN WO9414980-A.
XX
PD 07-JUL-1994.
XX
PF 20-DEC-1993; 93WO-US12388.
XX
PR 23-DEC-1992; 92US-0996783.
PR 17-SEP-1993; 93US-0123936.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
XX
DR WPI; 1994-234711/28.
XX
PT Sequence-directed DNA-binding molecules - useful in
PT pharmaceuticals and as molecular reagents
XX
PS Claim 28; Page 243; 587pp; English.
XX
CC A DNA protein-binding assay is provided, useful for screening
CC libraries of synthetic or biological cpds. for their ability
CC to bind DNA test sequences. The assay is versatile in that any

```
CC   number of test sequences can be tested by placing the test sequence
CC   adjacent to a defined protein-binding screening sequence.  Binding
CC   of mols. to these test sequences changes the binding characteristics
CC   of the protein mol. to its cognate binding sequence.  When such a mol.
CC   binds the test sequence, the equilibrium of the DNA:protein complexes
CC   is disturbed, generating changes in the concentration of free DNA probe.
CC   One application of this method is to eucaryotic general transcription
CC   factors (e.g. TFIID), where the target region is typically selected
CC   from DNA sequences adjacent to the binding site for the eucaryotic
CC   transcription factor.  Numerous exemplary test sequences are given:
CC   the sequences in AAQ69251-731 and AAQ69850 correspond to promoter
CC   targets (typically, TATA box-contg. sites) for human genes and the
CC   sequences in AAQ69732-849 correspond to promoter targets for viral genes.
CC    The test sequences may also be randomly generated.  DNA:protein
CC   interaction may be used for screening purposes, e.g. the Herpes Simplex
CC   Virus (HSV) origin of replication and UL9 (see AAQ69851-52, AAQ69865 and
CC   AAQ69891).
XX
SQ   Sequence 44 BP; 6 A; 15 C; 20 G; 3 T; 0 other;

  Query Match             1.0%;  Score 22.4;  DB 15;  Length 44;
  Best Local Similarity   72.5%;  Pred. No. 4.2e+04;
  Matches   29;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       9 CTGCGGCGGGTCCGGGCCCATGAGGCGACGAAGGAGGCGG 48
           || | ||||   |||||  |||||| | || ||| |||||
Db       4 CTCCCGCGGACGCGGGCATATGAGGAGGCGGAGGCGGCGG 43


RESULT 40
AAT63775
ID   AAT63775 standard; DNA; 44 BP.
XX
AC   AAT63775;
XX
DT   13-MAR-1997  (first entry)
XX
DE   Human Na/K ATPase alpha 3 subunit (ATP1 A3) gene TFIID binding site.
XX
KW   Duplex DNA; target region; binding characteristic; DNA binding protein;
KW   TFIID; transcription factor; binding site; inhibition; enhance;
KW   cancer; inherited genetic disorder; ds.
XX
OS   Homo sapiens.
XX
PN   US5578444-A.
XX
PD   26-NOV-1996.
XX
PF   27-JUN-1991;   91US-0723618.
XX
PR   20-DEC-1993;   93US-0171389.
PR   27-JUN-1991;   91US-0723618.
PR   23-DEC-1992;   92US-0996783.
PR   17-SEP-1993;   93US-0123936.
XX
PA   (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI   Andrews BM,  Cantor CR,  Edwards CA,  Fry KE,  Turin LM;
XX
DR   WPI; 1997-020402/02.
XX
PT   Altering binding characteristics of DNA binding proteins to duplex
PT   DNA - by attaching specific small cpd. to target region close to the
PT   protein's binding site, useful in treatment of viral disease, cancer
PT   etc
XX
PS   Claim 6; Column 131; 264pp; English.
XX
CC   The sequences given in AAT63713-4312 represent duplex DNA's which act
CC   as target regions in the method of the invention.  The method for
CC   altering the binding characteristics of a DNA-binding protein to duplex
CC   DNA comprises contacting the duplex DNA with a small molecule which
CC   binds sequence-specifically to a target region, where, when the small
CC   molecule is bound to the target region, it is adjacent to, but not
CC   overlapping by more than 4 bp, a binding site for a DNA-binding protein.
CC   The small molecule is added at a concentration effective to alter the
CC   binding of the DNA binding protein, pref. TFIID, to its binding site on
CC   the duplex DNA.  The binding of the small molecule may inhibit or
CC   enhance the binding of the DNA-binding protein to its binding site.  The
CC   compounds isolated using this method are potentially useful as
CC   therapeutic agents for treatment of any disease which involves a
CC   specific DNA sequence, e.g. cancer, or inherited genetic disorders etc.
CC   The method is suitable for screening large biological or chemical
CC   libraries and allows determination of sequence-specific and relative
CC   affinities of known DNA-binding agents for different DNA sequences.
CC   The design of these duplex DNA's allows a single DNA:protein interaction
CC   to be used for screening sequence-specific, or preferential, DNA binding
CC   proteins that recognise almost any possible sequence (see also AAT49539-
CC   74).
XX
SQ   Sequence 44 BP; 6 A; 15 C; 20 G; 3 T; 0 other;

  Query Match             1.0%;  Score 22.4;  DB 18;  Length 44;
  Best Local Similarity   72.5%;  Pred. No. 4.2e+04;
  Matches   29;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       9 CTGCGGCGGGTCCGGGCCCATGAGGCGACGAAGGAGGCGG 48
           || | ||||   |||||  |||||| | || ||| |||||
Db       4 CTCCCGCGGACGCGGGCATATGAGGAGGCGGAGGCGGCGG 43


RESULT 41
AAX17063
ID   AAX17063 standard; DNA; 44 BP.
XX
AC   AAX17063;
XX
DT   06-MAY-1999  (first entry)
XX
DE   Test sequence from human sodium/potassium ATPase alpha 3 subunit.
XX
KW   Test sequence; DNA-binding molecule; screening sequence; human;
KW   nucleic acid amplification; target; viral; ds.
XX
OS   Homo sapiens.
XX
PN   US5869241-A.
XX
PD   09-FEB-1999.
XX
PF   07-JUN-1995;   95US-0475228.
XX
PR   20-DEC-1993;   93US-0171389.
PR   27-JUN-1991;   91US-0723618.
PR   23-DEC-1992;   92US-0996783.
PR   17-SEP-1993;   93US-0123936.
PR   07-JUN-1995;   95US-0475228.
XX
PA   (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI   Andrews BM,  Cantor CR,  Edwards CA,  Fry KE,  Turin LM;
XX
DR   WPI; 1999-152755/13.
XX
PT   Determination of DNA sequence preference of a DNA-binding molecule -
PT   based on inhibition of binding of protein to oligonucleotide
PT   sequence attached to test sequence
XX
PS   Claim 3; Columns 131-132; 270pp; English.
XX
CC   Sequences AAX17001 to AAX17600 represent specifically claimed target
CC   test sequences that are used in the method of the invention of
CC   determining the DNA sequence preference of a DNA-binding molecule. The
CC   method comprises: (i) adding a test molecule and a DNA-binding protein to
CC   a mixture of duplex DNA test oligonucleotides, each of the test
```

```
CC   oligonucleotides having a test sequence adjacent to a screening sequence,
CC   where the screening sequence binds to the DNA-binding protein with a
CC   binding affinity that is independent of the DNA sequence of the test
CC   sequence, and where the mixture of duplex DNA test oligonucleotides
CC   includes several test sequences; (ii) incubating the test molecule, the
CC   mixture of duplex DNA test oligonucleotides and the DNA-binding protein
CC   for a time sufficient to permit binding of the test molecule to test
CC   sequences in the duplex DNA; (iii) separating unbound test
CC   oligonucleotides from test oligonucleotides bound to binding protein;
CC   (iv) amplifying the unbound test oligonucleotides; (v) repeating steps
CC   (ii) to (iv); (vi) isolating the amplified test oligonucleotides; and
CC   (vii) sequencing the isolated test oligonucleotides. Test sequences
CC   AAX17001-X17481 and AAX17600 correspond to promoter targets for human
CC   genes and test sequences AAX17482-X17599 correspond to promoter targets
CC   for viral genes.
XX
SQ   Sequence 44 BP; 6 A; 15 C; 20 G; 3 T; 0 other;

  Query Match             1.0%;  Score 22.4;  DB 20;  Length 44;
  Best Local Similarity   72.5%;  Pred. No. 4.2e+04;
  Matches   29;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       9 CTGCGGCGGGTCCGGGCCCATGAGGCGACGAAGGAGGCGG 48
           || | ||||   |||||  |||||| | || ||| |||||
Db       4 CTCCCGCGGACGCGGGCATATGAGGAGGCGGAGGCGGCGG 43


RESULT 42
ABK82554
ID   ABK82554 standard; DNA; 44 BP.
XX
AC   ABK82554;
XX
DT   27-AUG-2002  (first entry)
XX
DE   DNA binding molecule screening method test sequence #63.
XX
KW   DNA binding molecule screening; inhibition of transcription;
KW   infection; human immunodeficiency virus; HIV; parasite; cancer;
KW   cardiovascular; respiratory; gastrointestinal; endocrine; metabolic;
KW   rheumatic; immunological; haematological; neurological;
KW   psychiatric; dermatological; ophthalmological; musculo-skeletal;
KW   urogenital disorder; ss.
XX
OS   Synthetic.
XX
PN   US6384208-B1.
XX
PD   07-MAY-2002.
XX
PF   15-JUL-1999;   99US-0354947.
XX
PR   20-DEC-1993;   93US-0171389.
PR   07-JUN-1995;   95US-0482080.
PR   27-JUN-1991;   91US-0723618.
PR   23-DEC-1992;   92US-0996783.
PR   17-SEP-1993;   93US-0123936.
XX
PA   (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI   Edwards CA,  Cantor CR,  Andrews BM,  Turin LM,  Fry KE;
XX
DR   WPI; 2002-442819/47.
XX
PT   Decreasing transcriptional activity of genes for treating infections or
PT   cancer, by administration of an agent that binds to two non-overlapping
PT   regions of the gene  -
XX
PS   Example 15; SEQ ID No 63; 98pp; English.
XX
CC   The invention relates to a method of decreasing transcriptional activity
CC   in a duplex deoxyribonucleic acid (DNA) template (T1) comprising
CC   contacting (T1) with a binding agent comprising at least one small duplex
CC   DNA-binding molecule (T2) coupled to at least one other small duplex-
CC   binding molecule that binds to a non-overlapping region of target
CC   sequence (TS). The method is useful for inhibiting transcription of a
CC   range of disease-related genes for treating infections (by viruses,
CC   including human immunodeficiency virus, bacteria, fungi, protozoa
CC   and parasites), cancer, cardiovascular, respiratory, gastrointestinal,
CC   endocrine/metabolic, rheumatic/immunological, haematological,
CC   neurological, psychiatric, dermatological, ophthalmological,
CC   musculo-skeletal, genetic or urogenital disorders. The method provides
CC   sequence-specific inhibition of transcription of pathological genes
CC   without affecting transcription of cellular genes regulated by the same
CC   transcription factor, and can be applied to regulation of any gene.
CC   ABK82492-ABK83155 represent DNA binding molecule test sequences used in
CC   the method of the invention.
XX
SQ   Sequence 44 BP; 6 A; 15 C; 20 G; 3 T; 0 other;

  Query Match             1.0%;  Score 22.4;  DB 24;  Length 44;
  Best Local Similarity   72.5%;  Pred. No. 4.2e+04;
  Matches   29;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       9 CTGCGGCGGGTCCGGGCCCATGAGGCGACGAAGGAGGCGG 48
           || | ||||   |||||  |||||| | || ||| |||||
Db       4 CTCCCGCGGACGCGGGCATATGAGGAGGCGGAGGCGGCGG 43


RESULT 43
AAF29313/c
ID   AAF29313 standard; DNA; 48 BP.
XX
AC   AAF29313;
XX
DT   18-APR-2001  (first entry)
XX
DE   Primer base sequence used to illustrate primer selection method.
XX
KW   Primer; optimum sequence; differential display; ss.
XX
OS   Synthetic.
XX
PN   JP2000308487-A.
XX
PD   07-NOV-2000.
XX
PF   30-MAR-1999;   99JP-0088410.
XX
PR   30-MAR-1999;   99JP-0088410.
XX
PA   (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
DR   WPI; 2001-046077/06.
XX
PT   Selection of primer base for optimizing primer selection comprises
PT   obtaining an optimum sequence for differential display from an
PT   expression gene data base -
XX
PS   Disclosure; Fig 9; 13pp; Japanese.
XX
CC   This invention relates to a method for selecting the sequence of a
CC   primer. The method comprises obtaining an optimum sequence for
CC   differential display from an expression gene data base, and using the
CC   base sequences most frequently expressed as the primer candidates in the
CC   order of frequency. The optimum primer group characterised by the use of
CC   genetic algorithm from the primer candidates is selected. The method is
CC   used for selecting a primer sequence quickly. The present sequence
CC   represents a primer used in an illustration of the method of the
CC   invention.
XX
SQ   Sequence 48 BP; 15 A; 3 C; 3 G; 27 T; 0 other;

  Query Match             1.0%;  Score 22.4;  DB 22;  Length 48;
  Best Local Similarity   66.7%;  Pred. No. 4.4e+04;
  Matches   32;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;
```

```
Qy   1757 TTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATAGT 1804
          ||| |||||| | ||||   |  |||||||  ||| ||| | |  | |
Db     48 TTTGAAAAATAATTCAAATATATAAAAAAATTCTTCAAGAAGATAAAT 1
```

RESULT 44
AAF29318/c
ID AAF29318 standard; DNA; 48 BP.
XX
AC AAF29318;
XX
DT 18-APR-2001 (first entry)
XX
DE Primer base sequence used to illustrate primer selection method.
XX
KW Primer; optimum sequence; differential display; ss.
XX
OS Synthetic.
XX
PN JP2000308487-A.
XX
PD 07-NOV-2000.
XX
PF 30-MAR-1999; 99JP-0088410.
XX
PR 30-MAR-1999; 99JP-0088410.
XX
PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
DR WPI; 2001-046077/06.
XX
PT Selection of primer base for optimizing primer selection comprises
PT obtaining an optimum sequence for differential display from an
PT expression gene data base -
XX
PS Disclosure; Fig 9; 13pp; Japanese.
XX
CC This invention relates to a method for selecting the sequence of a
CC primer. The method comprises obtaining an optimum sequence for
CC differential display from an expression gene data base, and using the
CC base sequences most frequently expressed as the primer candidates in the
CC order of frequency. The optimum primer group characterised by the use of
CC genetic algorithm from the primer candidates is selected. The method is
CC used for selecting a primer sequence quickly. The present sequence
CC represents a primer used in an illustration of the method of the
CC invention.
XX
SQ Sequence 48 BP; 15 A; 3 C; 3 G; 27 T; 0 other;

```
  Query Match             1.0%;  Score 22.4;  DB 22;  Length 48;
  Best Local Similarity   66.7%;  Pred. No. 4.4e+04;
  Matches   32;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   1757 TTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATAGT 1804
          ||| |||||| | ||||   |  |||||||  ||| ||| | |  | |
Db     48 TTTGAAAAATAATTCAAATATATAAAAAAATTCTTCAAGAAGATAAAT 1
```

RESULT 45
AAL27843/c
ID AAL27843 standard; DNA; 51 BP.
XX
AC AAL27843;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #1051.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 1680; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 51 BP; 7 A; 5 C; 2 G; 37 T; 0 other;

```
  Query Match             1.0%;  Score 22.4;  DB 22;  Length 51;
  Best Local Similarity   66.7%;  Pred. No. 4.6e+04;
  Matches   32;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   1755 ATTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
          |||  | ||||  | | ||||  | |||||||||   |||  |||| |
Db     50 ATTAAATAAAAGAACTGAATGTGGTAAAAAAAAAAAAAAAAAAAAAA 3
```

RESULT 46
AAL32091
ID AAL32091 standard; DNA; 51 BP.
XX
AC AAL32091;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #5299.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;

KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 2912; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 51 BP; 18 A; 6 C; 5 G; 22 T; 0 other;

Query Match 1.0%; Score 22.4; DB 22; Length 51;
Best Local Similarity 66.7%; Pred. No. 4.6e+04;
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1592 GTATGAATCTATGACAGAGCAATGCTTTTAATGAATTTAAGGCAAAAA 1639
||| | ||| |||||||||| ||| || |||| || || ||
Db 4 GTAGATTTTTATTACAGAGCAATTTTTTCTATCAATTCAATGCTTTAA 51

RESULT 47
AAQ69462
ID AAQ69462 standard; DNA; 45 BP.
XX
AC AAQ69462;
XX
DT 27-FEB-1995 (first entry)
XX
DE Human interleukin 4 gene, target region.
XX
KW DNA protein-binding assay; test sequence; screening sequence;
KW promoter; target; TATA box; Herpes Simplex Virus; HSV;
KW origin of replication; UL9; transcription factor; TFIID: ds.
XX
OS Synthetic.
XX
PN WO9414980-A.
XX
PD 07-JUL-1994.
XX
PF 20-DEC-1993; 93WO-US12388.
XX
PR 23-DEC-1992; 92US-0996783.
PR 17-SEP-1993; 93US-0123936.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
XX
DR WPI; 1994-234711/28.
XX
PT Sequence-directed DNA-binding molecules - useful in
PT pharmaceuticals and as molecular reagents
XX
PS Claim 28; Page 318; 587pp; English.
XX
CC A DNA protein-binding assay is provided, useful for screening
CC libraries of synthetic or biological cpds. for their ability
CC to bind DNA test sequences. The assay is versatile in that any
CC number of test sequences can be tested by placing the test sequence
CC adjacent to a defined protein-binding screening sequence. Binding
CC of mols. to these test sequences changes the binding characteristics
CC of the protein mol. to its cognate binding sequence. When such a mol.
CC binds the test sequence, the equilibrium of the DNA:protein complexes
CC is disturbed, generating changes in the concentration of free DNA probe.
CC One application of this method is to eucaryotic general transcription
CC factors (e.g. TFIID), where the target region is typically selected
CC from DNA sequences adjacent to the binding site for the eucaryotic
CC transcription factor. Numerous exemplary test sequences are given:
CC the sequences in AAQ69251-731 and AAQ69850 correspond to promoter
CC targets (typically, TATA box-contg. sites) for human genes and the
CC sequences in AAQ69732-849 correspond to promoter targets for viral genes.
CC The test sequences may also be randomly generated. DNA:protein
CC interaction may be used for screening purposes, e.g. the Herpes Simplex
CC Virus (HSV) origin of replication and UL9 (see AAQ69851-52, AAQ69865 and
CC AAQ69891).
XX
SQ Sequence 45 BP; 13 A; 10 C; 7 G; 15 T; 0 other;

Query Match 0.9%; Score 22.2; DB 15; Length 45;
Best Local Similarity 77.1%; Pred. No. 4.8e+04;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 933 TGACAAAATCCTCAAATGTAAACTCAATTTGCCTC 967
| || ||| || |||||||||||| ||| ||||
Db 9 TAACGAAAATTTCCAATGTAAACTCATTTTCCCTC 43

RESULT 48
AAT63924
ID AAT63924 standard; DNA; 45 BP.
XX
AC AAT63924;
XX
DT 17-MAR-1997 (first entry)
XX
DE Human interleukin-4 gene TFIID binding site.
XX
KW Duplex DNA; target region; binding characteristic; DNA binding protein;
KW TFIID; transcription factor; binding site; inhibition; enhance; IL-2;
KW cancer; inherited genetic disorder; alpha-D-galactosidase A; ds.
XX
OS Homo sapiens.
XX
PN US5578444-A.
XX
PD 26-NOV-1996.
XX
PF 27-JUN-1991; 91US-0723618.
XX
PR 20-DEC-1993; 93US-0171389.
PR 27-JUN-1991; 91US-0723618.
PR 23-DEC-1992; 92US-0996783.

PR 17-SEP-1993; 93US-0123936.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
XX
DR WPI; 1997-020402/02.
XX
PT Altering binding characteristics of DNA binding proteins to duplex
PT DNA - by attaching specific small cpd. to target region close to the
PT protein's binding site, useful in treatment of viral disease, cancer
PT etc
XX
PS Claim 6; Column 207-208; 264pp; English.
XX
CC The sequences given in AAT63713-4312 represent duplex DNA's which act
CC as target regions in the method of the invention. The method for
CC altering the binding characteristics of a DNA-binding protein to duplex
CC DNA comprises contacting the duplex DNA with a small molecule which
CC binds sequence-specifically to a target region, where, when the small
CC molecule is bound to the target region, it is adjacent to, but not
CC overlapping by more than 4 bp, a binding site for a DNA-binding protein.
CC The small molecule is added at a concentration effective to alter the
CC binding of the DNA binding protein, pref. TFIID, to its binding site on
CC the duplex DNA. The binding of the small molecule may inhibit or
CC enhance the binding of the DNA-binding protein to its binding site. The
CC compounds isolated using this method are potentially useful as
CC therapeutic agents for treatment of any disease which involves a
CC specific DNA sequence, e.g. cancer, or inherited genetic disorders etc.
CC The method is suitable for screening large biological or chemical
CC libraries and allows determination of sequence-specific and relative
CC affinities of known DNA-binding agents for different DNA sequences.
CC The design of these duplex DNA's allows a single DNA:protein interaction
CC to be used for screening sequence-specific, or preferential, DNA binding
CC proteins that recognise almost any possible sequence (see also AAT49539-
CC 74).
XX
SQ Sequence 45 BP; 13 A; 10 C; 7 G; 15 T; 0 other;

Query Match 0.9%; Score 22.2; DB 18; Length 45;
Best Local Similarity 77.1%; Pred. No. 4.8e+04;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 933 TGACAAAATCCTCAAATGTAAACTCAATTTGCCTC 967
| || ||| || |||||||||||| ||| ||||
Db 9 TAACGAAAATTTCCAATGTAAACTCATTTTCCCTC 43

RESULT 49
AAX17212
ID AAX17212 standard; DNA; 45 BP.
XX
AC AAX17212;
XX
DT 06-MAY-1999 (first entry)
XX
DE Test sequence from human interleukin 4 gene.
XX
KW Test sequence; DNA-binding molecule; screening sequence; human;
KW nucleic acid amplification; target; viral; ds.
XX
OS Homo sapiens.
XX
PN US5869241-A.
XX
PD 09-FEB-1999.
XX
PF 07-JUN-1995; 95US-0475228.
XX
PR 20-DEC-1993; 93US-0171389.
PR 27-JUN-1991; 91US-0723618.
PR 23-DEC-1992; 92US-0996783.
PR 17-SEP-1993; 93US-0123936.
PR 07-JUN-1995; 95US-0475228.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
XX
DR WPI; 1999-152755/13.
XX
PT Determination of DNA sequence preference of a DNA-binding molecule -
PT based on inhibition of binding of protein to oligonucleotide
PT sequence attached to test sequence
XX
PS Claim 3; Columns 209-210; 270pp; English.
XX
CC Sequences AAX17001 to AAX17600 represent specifically claimed target
CC test sequences that are used in the method of the invention of
CC determining the DNA sequence preference of a DNA-binding molecule. The
CC method comprises: (i) adding a test molecule and a DNA-binding protein to
CC a mixture of duplex DNA test oligonucleotides, each of the test
CC oligonucleotides having a test sequence adjacent to a screening sequence,
CC where the screening sequence binds to the DNA-binding protein with a
CC binding affinity that is independent of the DNA sequence of the test
CC sequence, and where the mixture of duplex DNA test oligonucleotides
CC includes several test sequences; (ii) incubating the test molecule, the
CC mixture of duplex DNA test oligonucleotides and the DNA-binding protein
CC for a time sufficient to permit binding of the test molecule to test
CC sequences in the duplex DNA; (iii) separating unbound test
CC oligonucleotides from test oligonucleotides bound to binding protein;
CC (iv) amplifying the unbound test oligonucleotides; (v) repeating steps
CC (ii) to (iv); (vi) isolating the amplified test oligonucleotides; and
CC (vii) sequencing the isolated test oligonucleotides. Test sequences
CC AAX17001-X17481 and AAX17600 correspond to promoter targets for human
CC genes and test sequences AAX17482-X17599 correspond to promoter targets
CC for viral genes.
XX
SQ Sequence 45 BP; 13 A; 10 C; 7 G; 15 T; 0 other;

Query Match 0.9%; Score 22.2; DB 20; Length 45;
Best Local Similarity 77.1%; Pred. No. 4.8e+04;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 933 TGACAAAATCCTCAAATGTAAACTCAATTTGCCTC 967
| || ||| || |||||||||||| ||| ||||
Db 9 TAACGAAAATTTCCAATGTAAACTCATTTTCCCTC 43

RESULT 50
ABK82703
ID ABK82703 standard; DNA; 45 BP.
XX
AC ABK82703;
XX
DT 27-AUG-2002 (first entry)
XX
DE DNA binding molecule screening method test sequence #212.
XX
KW DNA binding molecule screening; inhibition of transcription;
KW infection; human immunodeficiency virus; HIV; parasite; cancer;
KW cardiovascular; respiratory; gastrointestinal; endocrine; metabolic;
KW rheumatic; immunological; haematological; neurological;
KW psychiatric; dermatological; ophthalmological; musculo-skeletal;
KW urogenital disorder; ss.
XX
OS Synthetic.
XX
PN US6384208-B1.
XX
PD 07-MAY-2002.
XX
PF 15-JUL-1999; 99US-0354947.
XX
PR 20-DEC-1993; 93US-0171389.
PR 07-JUN-1995; 95US-0482080.

PR 27-JUN-1991; 91US-0723618.
PR 23-DEC-1992; 92US-0996783.
PR 17-SEP-1993; 93US-0123936.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Edwards CA, Cantor CR, Andrews BM, Turin LM, Fry KE;
XX
DR WPI; 2002-442819/47.
XX
PT Decreasing transcriptional activity of genes for treating infections or
PT cancer, by administration of an agent that binds to two non-overlapping
PT regions of the gene -
XX
PS Example 15; SEQ ID No 212; 98pp; English.
XX
CC The invention relates to a method of decreasing transcriptional activity
CC in a duplex deoxyribonucleic acid (DNA) template (T1) comprising
CC contacting (T1) with a binding agent comprising at least one small duplex
CC DNA-binding molecule (T2) coupled to at least one other small duplex-
CC binding molecule that binds to a non-overlapping region of target
CC sequence (TS). The method is useful for inhibiting transcription of a
CC range of disease-related genes for treating infections (by viruses,
CC including human immunodeficiency virus, bacteria, fungi, protozoa
CC and parasites), cancer, cardiovascular, respiratory, gastrointestinal,
CC endocrine/metabolic, rheumatic/immunological, haematological,
CC neurological, psychiatric, dermatological, ophthalmological,
CC musculo-skeletal, genetic or urogenital disorders. The method provides
CC sequence-specific inhibition of transcription of pathological genes
CC without affecting transcription of cellular genes regulated by the same
CC transcription factor, and can be applied to regulation of any gene.
CC ABK82492-ABK83155 represent DNA binding molecule test sequences used in
CC the method of the invention.
XX
SQ Sequence 45 BP; 13 A; 10 C; 7 G; 15 T; 0 other;

Query Match 0.9%; Score 22.2; DB 24; Length 45;
Best Local Similarity 77.1%; Pred. No. 4.8e+04;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 933 TGACAAAATCCTCAAATGTAAACTCAATTTGCCTC 967
| || ||| || |||||||||||| ||| ||||
Db 9 TAACGAAAATTTCCAATGTAAACTCATTTTCCCTC 43

RESULT 51
AAF29294
ID AAF29294 standard; DNA; 48 BP.
XX
AC AAF29294;
XX
DT 18-APR-2001 (first entry)
XX
DE Primer base sequence used to illustrate primer selection method.
XX
KW Primer; optimum sequence; differential display; ss.
XX
OS Synthetic.
XX
PN JP2000308487-A.
XX
PD 07-NOV-2000.
XX
PF 30-MAR-1999; 99JP-0088410.
XX
PR 30-MAR-1999; 99JP-0088410.
XX
PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
DR WPI; 2001-046077/06.
XX
PT Selection of primer base for optimizing primer selection comprises
PT obtaining an optimum sequence for differential display from an
PT expression gene data base -
XX
PS Disclosure; Fig 9; 13pp; Japanese.
XX
CC This invention relates to a method for selecting the sequence of a
CC primer. The method comprises obtaining an optimum sequence for
CC differential display from an expression gene data base, and using the
CC base sequences most frequently expressed as the primer candidates in the
CC order of frequency. The optimum primer group characterised by the use of
CC genetic algorithm from the primer candidates is selected. The method is
CC used for selecting a primer sequence quickly. The present sequence
CC represents a primer used in an illustration of the method of the
CC invention.
XX
SQ Sequence 48 BP; 17 A; 3 C; 2 G; 26 T; 0 other;

Query Match 0.9%; Score 22.2; DB 22; Length 48;
Best Local Similarity 69.8%; Pred. No. 5e+04;
Matches 30; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 2097 AACAACCTGAATCTTTTTTTTATATAAATATATATTTTTCAAA 2139
| | | |||| |||||||||||||| || | |||| ||
Db 5 ATCTTCTTGAAGAATTTTTTTATATAAAAATTTTATTTTTCAA 47

RESULT 52
AAL32145
ID AAL32145 standard; DNA; 51 BP.
XX
AC AAL32145;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #5353.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 2927; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is

CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 51 BP; 15 A; 7 C; 2 G; 27 T; 0 other;

Query Match 0.9%; Score 22.2; DB 22; Length 51;
Best Local Similarity 69.8%; Pred. No. 5.1e+04;
Matches 30; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 2110 TTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
||| |||||||||| | | ||||||||| | | | | ||| |
Db 2 TTTCTTTTATATATACACATATTTTTCTCAAAACATAGTTATT 44

RESULT 53
ABN40059
ID ABN40059 standard; DNA; 60 BP.
XX
AC ABN40059;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:12807.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 12807; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 16 A; 10 C; 21 G; 13 T; 0 other;

Query Match 0.9%; Score 22.2; DB 24; Length 60;
Best Local Similarity 61.0%; Pred. No. 5.6e+04;
Matches 36; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 800 AATACATGGCCCCTGAAATCTTGATGAGAAGTGGCCACAATCGTGCTGTGGATTGGTGG 858
| ||||| || || || | ||| | | || |||| | | |||||| ||||
Db 2 ATTACATAGCTCCCGAGGTGCTGAGCAAGAAAGGGCACAGTTTCGAGGTGGATGTGTGG 60

RESULT 54
ABN45158/c
ID ABN45158 standard; DNA; 60 BP.
XX
AC ABN45158;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:17906.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 17906; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the

CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 9 A; 26 C; 1 G; 24 T; 0 other;

Query Match 0.9%; Score 22.2; DB 24; Length 60;
Best Local Similarity 61.0%; Pred. No. 5.6e+04;
Matches 36; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 309 TCGGGTACTTGGTAAAGGGGGCTATGGAAAGGTTTTTCAAGTACGAAAAGTAACAGGAG 367
| ||||| | || || ||||| | || |||| ||| | | | || | || ||
Db 60 TAGGGTAATAGGGAATGGGGGTAAGTGAGAGGTGAGAAAAGCAAGGAGAGATAAAGTAG 2

RESULT 55
AAH44211/c
ID AAH44211 standard; DNA; 33 BP.
XX
AC AAH44211;
XX
DT 20-SEP-2001 (first entry)
XX
DE p70 S6K mutagenesis PCR primer #2.
XX
KW Protein kinase; identification; hydrophobic pocket; interacting;
KW cancer; diabetes; inhibition; apoptosis; tissue injury;
KW ischaemic injury; stroke; PCR primer; mutagenesis; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200144497-A2.
XX
PD 21-JUN-2001.
XX
PF 04-DEC-2000; 2000WO-GB04598.
XX
PR 02-DEC-1999; 99US-0168559.
XX
PA (UYDU-) UNIV DUNDEE.
XX
PI Alessi D, Biondi R;
XX
DR WPI; 2001-390252/41.
XX
PT Identifying modulators of protein kinase (PK) activity, useful in
PT developing drugs for treating cancer or diabetes, by measuring the
PT ability of the compound to modulate or mimic the interaction of PK with
PT interacting polypeptides -
XX
PS Example 1; Page 72; 180pp; English.
XX
CC The present invention describes a method for identifying a compound that
CC modulates protein kinase activity. The method comprises measuring the
CC ability of the compound to inhibit, promote or mimic the interaction of
CC a hydrophobic pocket-containing protein kinase with an interacting
CC polypeptide. The interacting polypeptide interacts with the hydrophobic
CC pocket of the protein kinase and/or comprises the amino acid sequence
CC Phe/Tyr-Xaa-Xaa-Phe/Tyr (I). The method is useful in screening assays
CC for developing pharmaceutical compounds or drugs. Compounds, polypeptides
CC or polynucleotides from the present invention are useful in medicine,
CC particularly in the manufacture of a medicament for treating a patient
CC in need of modulation of signalling by a hydrophobic pocket-containing
CC protein kinase. Specifically, the patient has cancer or diabetes or is
CC in need of inhibition of apoptosis, e.g. a patient suffering from tissue
CC injury or ischaemic injury, including stroke. The compound or
CC composition is also useful for inhibiting the degree or rate of
CC phosphorylation by the protein kinase. The interacting polypeptide or
CC compound is useful in methods of stabilising a hydrophobic pocket-
CC containing protein kinase, where the protein kinase is exposed to the
CC compound or polypeptide. AAB99786 to AAB99847 represent amino acid
CC sequences, and AAH44210 and AAH44211 represent oligonucleotide sequences,
CC used in the exemplification of the present invention.
XX
SQ Sequence 33 BP; 8 A; 10 C; 9 G; 6 T; 0 other;

Query Match 0.9%; Score 22; DB 22; Length 33;
Best Local Similarity 100.0%; Pred. No. 4.7e+04;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1269 GGCTCCATCTGTACTTGAAAGT 1290
||||||||||||||||||||||
Db 33 GGCTCCATCTGTACTTGAAAGT 12

RESULT 56
ABA96305/c
ID ABA96305 standard; DNA; 39 BP.
XX
AC ABA96305;
XX
DT 18-MAR-2002 (first entry)
XX
DE LDL receptor allele-specific probe GC1s B.
XX
KW Amplification product; pharmacogenetic testing; forensic; virus; SNP;
KW anthropology; paternity testing; pathogen; bacteria; yeast; HIV;
KW single nucleotide polymorphism; mini-sequencing; mutation;
KW triplet repeat expansion; chromosome rearrangement; cystic fibrosis;
KW sickle-cell anaemia; LDL receptor; probe; ss.
XX
OS Synthetic.
XX
PN WO200183823-A1.
XX
PD 08-NOV-2001.
XX
PF 30-APR-2001; 2001WO-US13979.
XX
PR 28-APR-2000; 2000US-200635P.
XX
PA (QUAN-) QUANTUM DOT CORP.
XX
PI Lai JH, Phillips VE, Watson AR;
XX
DR WPI; 2002-114152/15.
XX
PT Analysis of polynucleotides in a sample using generic capture sequences
PT comprises amplifying target polynucleotide, and utilizing the product
PT to indirectly assay the sample for the polynucleotide -
XX
PS Example 2; Page 47; 85pp; English.
XX
CC The invention relates to assaying for an amplification product from a
CC target polynucleotide. An amplification reaction is used to produce the
CC amplification product from the target polynucleotide so that it can be
CC used to indirectly assay the sample for the target polynucleotide. A
CC sample suspected of containing the target polynucleotide is contacted
CC with first and second primers to amplify the target polynucleotide; the
CC first primer comprises a tag sequence, the complement of which is formed
CC on the opposite strand during amplification and is referred to as a
CC capture label. A capture probe is provided that is conjugated to a
CC substrate and can bind to the capture sequence to form an amplification
CC product detection complex. Identification of the label in association
CC with the substrate demonstrates that the amplification product was formed
CC the target polynucleotide was present in the sample. The method is useful
CC for pharmacogenetic testing, for forensic or anthropological setting to
CC identify a species or individual which was the source of a specimen.

```
CC   The method can also be used for paternity testing, testing for
CC   compatibility of prospective tissue or blood donors and in screening for
CC   heredity disorders. Other applications include gene expression studies,
CC   human population genetics, to detect contaminants or pathogens including
CC   bacteria, yeast, viruses, for HIV subtyping and to detect single
CC   nucleotide polymorphisms associated with particular alleles or subsets of
CC   allele. The method is also useful for mini-sequencing and for detection
CC   of mutations, including without limitation SNPs, insertions, deletions,
CC   transitions, transversions, inversions, frame shifts, triplet repeat
CC   expansions and chromosome rearrangements. The method is useful to detect
CC   nucleotide sequences associated with increased risk of diseases or
CC   disorders, including cystic fibrosis, Tay-Sachs and sickle-cell anaemia.
CC   The present sequence is that of a probe for detection of allele-specific
CC   genomic DNA sequences from the LDL receptor in the method of the
CC   invention.
XX
SQ   Sequence 39 BP; 3 A; 5 C; 5 G; 26 T; 0 other;

  Query Match             0.9%;  Score 22;  DB 24;  Length 39;
  Best Local Similarity   73.7%;  Pred. No. 5.1e+04;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1763 AAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAA 1800
           |||||| ||| |||  || |||||||   |||  ||||
Db     39 AAATCAGTCAGTGGCCCAGAAAAAAAAAAAAAAAAAAA 2


RESULT 57
AAX76470
ID   AAX76470 standard; DNA; 42 BP.
XX
AC   AAX76470;
XX
DT   05-AUG-1999  (first entry)
XX
DE   Human BRCA1 interacting protein gene B112 PCR primer B112.2J.
XX
KW   Human; BRCA1 interacting protein; B112; CtIP; tumour suppressor;
KW   cancer; therapy; PCR primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9927075-A1.
XX
PD   03-JUN-1999.
XX
PF   20-NOV-1998;   98WO-US24831.
XX
PR   21-NOV-1997;   97US-0975703.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
XX
PI   Bartel PL,  Tavtigian SV,  Teng DHF,  Wong AKC;
XX
DR   WPI; 1999-357827/30.
XX
PT   A carboxy-terminal BRCA1 interacting protein
XX
PS   Example 4; Page 42; 93pp; English.
XX
CC   The present invention describes a human BRCA1 interacting protein,
CC   designated B112. BRCA1 is a tumour suppressor protein. Methods and
CC   compositions from the present invention are useful for diagnosis of,
CC   determining predisposition to or lack of predisposition to, and
CC   treatment of human cancer, such as breast or pancreatic cancer, as a
CC   result of a mutation in CtIP or BRCA1. The methods and compositions can
CC   also be used in rational drug design for cancer therapeutics. The
CC   present sequence represents a PCR primer for B112 which is used in an
CC   example from the present invention.
XX
SQ   Sequence 42 BP; 15 A; 11 C; 10 G; 6 T; 0 other;

  Query Match             0.9%;  Score 22;  DB 20;  Length 42;
  Best Local Similarity   73.7%;  Pred. No. 5.3e+04;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   2034 TTTTCCCAGGCAAAATGGGAAGGCAAAGACAAAAGAAA 2071
           ||||||||| ||  | ||||||| ||  |||| | | |
Db      2 TTTTCCCAGTCACGACGGGAAGGAAATCACAAGAAACA 39


RESULT 58
AAZ67996
ID   AAZ67996 standard; DNA; 47 BP.
XX
AC   AAZ67996;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human map-related biallelic marker SEQ ID NO:2343.
XX
KW   Human genome; biallelic marker; high density disequilibrium map;
KW   genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW   haplotyping; hybridisation; identification; characterisation;
KW   diagnosis; single nucleotide polymorphism; SNP; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(24,T)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO9954500-A2.
XX
PD   28-OCT-1999.
XX
PF   21-APR-1999;   99WO-IB00822.
XX
PR   21-APR-1998;   98US-0082614.
PR   23-NOV-1998;   98US-0109732.
XX
PA   (GEST ) GENSET.
XX
PI   Cohen D,  Blumenfeld M,  Chumakov I;
XX
DR   WPI; 2000-013267/01.
XX
PT   Novel biallelic markers used to construct a high density disequilibrium
PT   map of the human genome  -
XX
PS   Claim 3; Page 730; 2745pp; English.
XX
CC   AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC   invention, which contain a polymorphic base at position 24 of their
CC   nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC   primers for the biallelic markers. The biallelic markers of the
CC   invention have a variety of uses: they can be used for high density
CC   mapping of the human genome, and in complex association studies and
CC   haplotyping studies which are useful in determining the genetic basis
CC   for disease states. Compositions and methods of the invention can also
CC   be useful for the identification of the targets for the development of
CC   pharmaceutical agents and diagnostic methods, as well as the
CC   characterisation of the differential efficacious responses to and side
CC   effects from pharmaceutical agents acting on a disease as well as other
CC   treatment.
CC   N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297
CC   and 3367, are not actually given a sequence in the Sequence Listing
CC   from the present invention.
XX
SQ   Sequence 47 BP; 33 A; 4 C; 1 G; 9 T; 0 other;

  Query Match             0.9%;  Score 22;  DB 21;  Length 47;
  Best Local Similarity   73.7%;  Pred. No. 5.6e+04;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;
```

```
Qy    1760 AAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCA 1797
           ||||||| ||| |||    |||| ||||| | ||| ||
Db       5 AAAAAATAAATAAATAAAACAAATAAAAAATAAAAACA 42


RESULT 59
AAA64611
ID   AAA64611 standard; DNA; 47 BP.
XX
AC   AAA64611;
XX
DT   02-JAN-2001  (first entry)
XX
DE   AU rich sequence in 3' UTR of human GM-CSF.
XX
KW   OLD-35; OLD-64; OLD-137; OLD-139; OLD-142; OLD-175; cancerous phenotype;
KW   cellular senescence; terminal differentiation; growth suppression;
KW   aging process; type I interferon; cancer cell; tissue regeneration; ss.
XX
OS   Homo sapiens.
XX
PN   WO200046391-A2.
XX
PD   10-AUG-2000.
XX
PF   02-FEB-2000; 2000WO-US02920.
XX
PR   02-FEB-1999;   99US-0243277.
XX
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI   Fisher PB,  Leszcyniecka M;
XX
DR   WPI; 2000-532905/48.
XX
PT   Novel isolated nucleic acid encoding an OLD-35 or OLD-64 protein useful
PT   in the treatment and detection of e.g. cancer and diseases involving
PT   cellular senescence -
XX
PS   Disclosure; Fig 7; 115pp; English.
XX
CC   The specification describes OLD-35, OLD-64, OLD-137, OLD-139, OLD-142
CC   and OLD-175 proteins. The OLD nucleic acids are useful for reversing
CC   the cancerous phenotype of a cancer cell, determining if a cell is
CC   senescent, growth arrested or terminally differentiated. They are also
CC   useful for reversing the aging process in a cell and degrading specific
CC   RNAs in a cell. The genes may also be used as a diagnostic indicator of
CC   cellular senescence, terminal differentiation and/or growth suppression
CC   and as a marker to identify drugs or small molecules that will induce
CC   or inhibit cellular senescence or terminal differentiation and type I
CC   interferons. The combination of Old-35 with other interacting proteins
CC   is useful for targeting the differentiation of specific cells. Old-35
CC   can be used to selectively stabilize specific mRNAs containing adenoviral
CC   rich 3' UTRs. The OLD proteins are useful for reversing the cancerous
CC   phenotype of a cancer cell and inhibiting the growth of a cancer cell.
CC   They are also useful for regenerating tissue. The present sequence
CC   represents an AU rich region from the 3' untranslated region (UTR)
CC   of human granulocyte-monocyte colony stimulating factor (GM-CSF).
XX
SQ   Sequence 47 BP; 18 A; 0 C; 0 G; 29 U; 0 other;

  Query Match             0.9%;  Score 22;  DB 21;  Length 47;
  Best Local Similarity  23.7%;  Pred. No. 5.6e+04;
  Matches    9;  Conservative   19;  Mismatches   10;  Indels    0;  Gaps    0;

Qy    2115 TTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
           : : :|:| |: :|:|::::: |||:|  : :::  ::
Db       4 UAUUUAUAUAUUUAUAUUUUUAAAAUAUUUAUUUAUUU 41


RESULT 60
AAF29302
ID   AAF29302 standard; DNA; 48 BP.
XX
AC   AAF29302;
XX
DT   18-APR-2001  (first entry)
XX
DE   Primer base sequence used to illustrate primer selection method.
XX
KW   Primer; optimum sequence; differential display; ss.
XX
OS   Synthetic.
XX
PN   JP2000308487-A.
XX
PD   07-NOV-2000.
XX
PF   30-MAR-1999;   99JP-0088410.
XX
PR   30-MAR-1999;   99JP-0088410.
XX
PA   (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
DR   WPI; 2001-046077/06.
XX
PT   Selection of primer base for optimizing primer selection comprises
PT   obtaining an optimum sequence for differential display from an
PT   expression gene data base -
XX
PS   Disclosure; Fig 9; 13pp; Japanese.
XX
CC   This invention relates to a method for selecting the sequence of a
CC   primer. The method comprises obtaining an optimum sequence for
CC   differential display from an expression gene data base, and using the
CC   base sequences most frequently expressed as the primer candidates in the
CC   order of frequency. The optimum primer group characterised by the use of
CC   genetic algorithm from the primer candidates is selected. The method is
CC   used for selecting a primer sequence quickly. The present sequence
CC   represents a primer used in an illustration of the method of the
CC   invention.
XX
SQ   Sequence 48 BP; 19 A; 3 C; 3 G; 23 T; 0 other;

  Query Match             0.9%;  Score 22;  DB 22;  Length 48;
  Best Local Similarity  73.7%;  Pred. No. 5.6e+04;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy    2102 CCTGAATCTTTTTTTTATATAAATATATATTTTTCAAA 2139
           | ||||   ||||||||||||||| || |  ||||  ||
Db      10 CTTGAAGAATTTTTTTATATAAAAATTTTATTTTTCAA 47


RESULT 61
ABL00059/c
ID   ABL00059 standard; DNA; 50 BP.
XX
AC   ABL00059;
XX
DT   05-MAR-2002  (first entry)
XX
DE   Human silent noncoding SNP oligonucleotide SEQ ID NO:50.
XX
KW   Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;
KW   immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;
KW   autoimmune disease; inflammation; cancer; nervous system disease;
KW   infection; polymorphic protein; ds.
XX
OS   Homo sapiens.
XX
PN   WO200138586-A2.
XX
PD   31-MAY-2001.
XX
PF   22-NOV-2000; 2000WO-US32311.
```

XX
PR 24-NOV-1999; 99US-0167383.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-355949/37.
XX
PT Isolated human nucleic acids comprising one or more single nucleotide
PT polymorphisms, useful for treating a subject suffering from a
PT pathology, e.g. autoimmune diseases, ascribed to the presence of a
PT sequence polymorphism -
XX
PS Claim 1; Page 260; 674pp; English.
XX
CC ABL00010 to ABL01104 represent human nucleic acid oligonucleotides
CC comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
CC to ABB56903 represent human peptides encoded by some of the SNP
CC oligonucleotides. The sequences from the present invention can have
CC immunosuppressive, cytostatic, antiinflammatory, neuroprotective and
CC antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC and antibodies from the present invention can be used for treating a
CC subject suffering from, at risk for, or suspected of, suffering from a
CC pathology ascribed to the presence of a sequence polymorphism. The
CC pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC the nervous system, and infection by pathogenic microorganisms. The SNPs
CC are also useful for determining which forms of a characterised
CC polymorphism are present in individuals. The antibodies may be used in
CC the detection, quantitation and/or cellular or tissue localisation of a
CC polymorphic protein (e.g., for use in measuring levels of the
CC polymorphic protein within appropriate physiological samples).
XX
SQ Sequence 50 BP; 16 A; 4 C; 8 G; 22 T; 0 other;

Query Match 0.9%; Score 22; DB 23; Length 50;
Best Local Similarity 67.4%; Pred. No. 5.7e+04;
Matches 31; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 2096 AAACAACCTGAATCTTTTTTTTATATAAATATATATTTTTCAAATA 2141
||||||||||| | ||| | ||||| ||||||||||||
Db 48 AAACAATCTTTTACACACTTTGGAAAAAATAGGTATTTTTCAAATA 3

RESULT 62
AAL30100/c
ID AAL30100 standard; DNA; 51 BP.
XX
AC AAL30100;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #3308.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 2335; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 51 BP; 10 A; 0 C; 4 G; 37 T; 0 other;

Query Match 0.9%; Score 22; DB 22; Length 51;
Best Local Similarity 67.4%; Pred. No. 5.8e+04;
Matches 31; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1757 TTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
|| ||||||||| || | ||||||||| ||| |||| |
Db 46 TTAAAAAAATCATACATTCCAAAAAAAAAAAAAAAAAAAAAAAAA 1

RESULT 63
AAL31018/c
ID AAL31018 standard; DNA; 51 BP.
XX
AC AAL31018;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #4226.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX

```
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 2600; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 5 A; 6 C; 2 G; 38 T; 0 other;

  Query Match             0.9%;  Score 22;  DB 22;  Length 51;
  Best Local Similarity   67.4%;  Pred. No. 5.8e+04;
  Matches   31;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   1743 AAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAA 1788
           ||||||| ||||  |  ||||||  ||  ||      |||||||||
Db     49 AAATAAAAGTGGCCTGGAAAAAAGAAAAAAAAAAAAAAAAAAAAA 4


RESULT 64
AAL31438/c
ID   AAL31438 standard; DNA; 51 BP.
XX
AC   AAL31438;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #4646.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;  99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 2722; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 10 A; 4 C; 2 G; 35 T; 0 other;

  Query Match             0.9%;  Score 22;  DB 22;  Length 51;
  Best Local Similarity   67.4%;  Pred. No. 5.8e+04;
  Matches   31;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   1757 TTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
           ||| || || ||||     ||| |||||||||   |||  |||| |
Db     49 TTTTAATAAACAATTTGCAGTGAAAAAAAAAAAAAAAAAAAAAAA 4


RESULT 65
AAH26598
ID   AAH26598 standard; mRNA; 51 BP.
XX
AC   AAH26598;
XX
DT   12-NOV-2001  (first entry)
XX
DE   Human GM-CSF gene 3' UTR AU-rich element.
XX
KW   Granulocyte-macrophage colony stimulating factor; GM-CSF;
KW   human; AU-rich element; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    5..9
FT                   /*tag=  a
FT                   /note= "AUUUA motif"
FT   misc_feature    9..13
FT                   /*tag=  b
FT                   /note= "AUUUA motif"
FT   misc_feature    13..17
FT                   /*tag=  c
FT                   /note= "AUUUA motif"
FT   misc_feature    30..34
FT                   /*tag=  d
FT                   /note= "AUUUA motif"
FT   misc_feature    34..38
FT                   /*tag=  e
FT                   /note= "AUUUA motif"
FT   misc_feature    38..42
FT                   /*tag=  f
FT                   /note= "AUUUA motif"
FT   misc_feature    42..46
FT                   /*tag=  g
FT                   /note= "AUUUA motif"
FT   misc_feature    46..50
```

```
FT                  /*tag=  h
FT                  /note= "AUUUA motif"
XX
PN   WO200164921-A1.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-US06782.
XX
PR   29-FEB-2000; 2000US-0515369.
XX
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI   Fisher PB,  Madireddi MT;
XX
DR   WPI; 2001-565508/63.
XX
PT   Melanoma differentiation associated gene-7 promoter capable of
PT   treating cancer comprises directing transcription of heterologous
PT   coding sequence encoding tumour suppressor polypeptide positioned
PT   downstream, useful for treating cancer  -
XX
PS   Disclosure; Fig 2C; 132pp; English.
XX
CC   The present sequence is that of an AU-rich sequence in the 3'
CC   untranslated region (3'UTR) of human granulocyte-macrophage
CC   colony stimulating factor mRNA.  The presence of AU-rich elements
CC   (AREs) in eukaryotic mRNAs correlates with rapid mRNA turnover
CC   and post-translational control.  The ARE consists of multiple AUUUA
CC   motifs or sequences resembling it.  A similar ARE sequence is found
CC   in the 3' UTR of the human melanoma differentiation associated
CC   gene-7 (Mda-7) gene (see AAH26596).  The invention provides
CC   recombinant expression constructs in which the human Mda-7 promoter
CC   (see AAH26595) is operably linked to a coding sequence encoding a
CC   tumour suppressor protein.  A pharmaceutical composition including
CC   the recombinant expression construct is used in a claimed method of
CC   treating melanoma, neuroblastoma, astrocytoma, glioblastoma
CC   multiforme, cervical cancer, breast cancer, colon cancer, prostate
CC   cancer, osteosarcoma, chondrosarcoma or a cancer of the central
CC   nervous system.
XX
SQ   Sequence 51 BP; 19 A; 0 C; 0 G; 32 U; 0 other;

  Query Match             0.9%;  Score 22;  DB 22;  Length 51;
  Best Local Similarity   23.7%;  Pred. No. 5.8e+04;
  Matches    9;  Conservative   19;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   2115 TTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
           : : :|:| |: :|:|::::: |||:|  : :::  ::
Db      4 UAUUUAUAUAUUUAUAUUUUUAAAAUAUUUAUUUAUUU 41


RESULT 66
AAS19341/c
ID   AAS19341 standard; DNA; 58 BP.
XX
AC   AAS19341;
XX
DT   20-MAR-2002  (first entry)
XX
DE   Oligonucleotide 5982 used to contruct plasmid XL2725.
XX
KW   ss; DNA purification; triple helix; plasmid purification;
KW   oligonucleotide 5982; XL2725.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   repeat_region   5..55
FT                   /*tag=  a
FT                   /rpt_type= "TANDEM"
FT   repeat_unit     5..7
FT                   /*tag=  b
FT                   /note= "CCT repeat type"
XX
PN   WO200192511-A2.
XX
PD   06-DEC-2001.
XX
PF   25-MAY-2001; 2001WO-US17122.
XX
PR   26-MAY-2000; 2000US-0580923.
XX
PA   (AVET ) AVENTIS PHARMA SA.
XX
PI   Crouzet J,  Scherman D,  Wils P,  Blanche F,  Cameron B;
XX
DR   WPI; 2002-097772/13.
XX
PT   Purifying double-stranded (ds) DNA from a solution containing dsDNA and
PT   other components, comprises passing the solution through a support
PT   comprising a covalently coupled oligonucleotide able to form a triple
PT   helix with the dsDNA  -
XX
PS   Example 7.1; Page 20; 40pp; English.
XX
CC   This invention comprises a method of purifying double-stranded DNA from
CC   a solution containing the double-stranded DNA mixed with other
CC   components, comprising passing the solution through a support comprising
CC   a covalently coupled oligonucleotide capable of forming a triple helix
CC   with the double-stranded DNA by hybridisation with a specific sequence
CC   present in the double-stranded DNA. The method is useful for purifying
CC   double-stranded DNA contained in a solution and mixed with other
CC   components. The new method is a simple, rapid and effective method for
CC   DNA purification, and makes it possible to obtain especially high
CC   purities with high yields. The method enables DNA to be purified from
CC   complex mixtures comprising other nucleic acids, proteins, endotoxins,
CC   nucleases and the like. The supports may be readily recycled, and the
CC   DNAs obtained display improved properties to pharmaceutical safety.
CC   Further, the method entails only one step contrary to prior art.
CC   The present sequence represents an oligonucleotide 5982 used to
CC   create the XL2725 plasmid which was used in an example of the DNA
CC   purification method of the invention.
XX
SQ   Sequence 58 BP; 2 A; 36 C; 1 G; 19 T; 0 other;

  Query Match             0.9%;  Score 22;  DB 24;  Length 58;
  Best Local Similarity   63.0%;  Pred. No. 6.2e+04;
  Matches   34;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy    123 GGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAGTT 176
          |||  ||   ||||| | ||     ||||| |||  |||||||| ||   | ||
Db     54 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAATT 1


RESULT 67
AAV04153
ID   AAV04153 standard; DNA; 60 BP.
XX
AC   AAV04153;
XX
DT   22-JUN-1998  (first entry)
XX
DE   Blood group antigen binding (Bab) adhesin DNA fragment.
XX
KW   Blood group antigen binding adhesin; Baba; Babb; infection;
KW   gastritis; acid peptic disease; therapy; diagnosis; vaccine;
KW   immunisation; ss.
XX
OS   Helicobacter pylori.
XX
PN   WO9747646-A1.
XX
PD   18-DEC-1997.
XX
PF   10-JUN-1997;   97WO-SE01009.
```

```
XX
PR   19-MAR-1997;   97SE-0001014.
PR   10-JUN-1996;   96SE-0002287.
XX
PA   (BORE/) BOREN T.
XX
PI   Arnqvist A,  Boren T,  Hammarstrom L,  Ilver D,  Normark S;
PI   Marstroem L;
XX
DR   WPI; 1998-052240/05.
XX
PT   Helicobacter pylori blood group antigen binding adhesin protein -
PT   that binds fucosylated blood group antigen, useful to diagnose,
PT   prevent and treat H. pylori infection
XX
PS   Claim 11; Page 26; 53pp; English.
XX
CC   This DNA fragment comprises a 5' fragment of DNA coding for novel
CC   Helicobacter pylori blood group antigen binding (Bab) adhesin
CC   proteins that bind specifically to fucosylated blood group
CC   antigens.  Examples include the claimed baba (see AAV04154) and
CC   babb (see AAV04155) genes of H. pylori strain CCUG 17875.  The bab
CC   adhesins (see also AAW41522 and AAW41523) and DNA can be used in the
CC   diagnosis, treatment and/or prophylaxis of H. pylori induced
CC   infections such as gastritis and acid peptic disease, e.g. by
CC   active immunisation.  The genes can also be used in the recombinant
CC   production of Bab adhesin polypeptides.
XX
SQ   Sequence 60 BP; 19 A; 13 C; 16 G; 12 T; 0 other;

  Query Match             0.9%;  Score 22;  DB 19;  Length 60;
  Best Local Similarity   63.0%;  Pred. No. 6.3e+04;
  Matches   34;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy      49 GACGGCTTTTACCCAGCCCCGGACTTCCGAGACAGGGAAGCTGAGGACATGGCA 102
           ||||||||||||| ||  |   | ||  | |  | | ||||| |   | |||| |
Db       7 GACGGCTTTTACACAAGCGTAGGCTATCAAATCGGTGAAGCCGCTCAAATGGTA 60


RESULT 68
AAH93339
ID   AAH93339 standard; DNA; 60 BP.
XX
AC   AAH93339;
XX
DT   04-OCT-2001  (first entry)
XX
DE   Plasmodium falciparum MAL3PC polynucleotide SEQ ID NO 60.
XX
KW   Human; antisense-therapy; gene-therapy; diagnostic; forensic;
KW   gene mapping; ds.
XX
OS   Plasmodium falciparum.
XX
PN   WO200152616-A2.
XX
PD   26-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-US35190.
XX
PR   23-DEC-1999;   99US-0471275.
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-451890/48.
XX
PT   Isolated polypeptide for treatment of diseases, diagnostics, raising
PT   antibodies and research use -
XX
PS   Example 4; Page 122; 135pp; English.
XX
CC   The invention relates to an isolated human polynucleotide (AAH75398)
CC   encoding a novel polypeptide (AAG64527) useful in antisense-therapy and
CC   gene-therapy, in diagnostics, forensics, gene mapping and identification
CC   of mutations responsible for genetic disorders and other traits.
CC   Polynucleotide sequences with potential homology were also identified
CC   (AAH93283-AAH93356).
XX
SQ   Sequence 60 BP; 31 A; 4 C; 1 G; 24 T; 0 other;

  Query Match             0.9%;  Score 22;  DB 22;  Length 60;
  Best Local Similarity   67.4%;  Pred. No. 6.3e+04;
  Matches   31;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy    2096 AAACAACCTGAATCTTTTTTTTATATAAATATATATTTTTCAAATA 2141
           ||||||| |        | | |||||| |||||||| ||  |||||
Db       2 AAACAACATATGATACATATATATATATATATATATATTATAAATA 47


RESULT 69
ABN38032
ID   ABN38032 standard; DNA; 60 BP.
XX
AC   ABN38032;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:10780.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 10780; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
```

CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 18 A; 14 C; 17 G; 11 T; 0 other;

Query Match 0.9%; Score 22; DB 24; Length 60;
Best Local Similarity 63.0%; Pred. No. 6.3e+04;
Matches 34; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 2032 ACTTTTCCCAGGCAAAATGGGAAGGCAAAGACAAAAGAAACTTACCAATTGATG 2085
|| || |||||| |||| | | |||||| | || |||| |||||
Db 1 ACCTTGTACAGGCAGCATGGAGAAACTAAGACAGAGTGTCCTGCCCAAGTGATG 54

RESULT 70
AAT61581/c
ID AAT61581 standard; DNA; 50 BP.
XX
AC AAT61581;
XX
DT 28-OCT-1997 (first entry)
XX
DE VH and scFv antibody library VH back-primer VH 3,5.
XX
KW Human; monoclonal antitumour antibody; peripheral blood lymphocyte;
KW cancer; tumourigenesis; anticancer vaccine; PCR;
KW polymerase chain reaction; ss.
XX
OS Synthetic.
XX
PN WO9702479-A2.
XX
PD 23-JAN-1997.
XX
PF 28-JUN-1996; 96WO-IB01032.
XX
PR 30-JUN-1995; 95US-0497647.
XX
PA (UYYA) UNIV YALE.
XX
PI Cai X, Garen A;
XX
DR WPI; 1997-109061/10.
XX
PT Prodn. of human monoclonal anti-tumour antibodies - by screening a
PT fusion phage library produced using peripheral blood lymphocytes
PT from a cancer patient
XX
PS Example 3; Page 39; 82pp; English.
XX
CC A process for isolating and synthesising human monoclonal anti-tumour
CC antibodies has been produced. The process involves: (a) constructing at
CC least one fusion phage library from the peripheral blood lymphocytes
CC (PBLs) of a cancer patient; (b) screening for anti-tumour antibodies in
CC the phage library in a binding assay with cultured tumour cells of the
CC same type as the patient's tumour; (c) removing extraneous antibodies by
CC absorption against normal human cells; (d) cloning the phage selected in
CC step (b) and (c); (e) assaying the specificity of the cloned phage by
CC incubating the phage with at least two types of cultured normal cells;
CC and (f) further testing the specificity of cloned phage that do not bind
CC to either cell line of cultured normal cells in further binding assays
CC to cultured tumour cells derived from more than one other tumour that is
CC not the patient's tumour. The present sequence represents a VH back-
CC primer involved in the construction of VH and scFv libraries. The
CC antibodies produced can be used for diagnostic and therapeutic
CC applications and for isolating tumour antigens for studying
CC tumourigenesis or for use as anti-cancer vaccines. The human antibodies
CC have low immunogenicity in humans compared to murine monoclonal
CC antibodies (MAbs). Since the antibodies are isolated from fusion phage
CC libraries, their affinity and specificity for a tumour cell line can be
CC improved by genetic manipulations.
XX
SQ Sequence 50 BP; 6 A; 11 C; 20 G; 10 T; 3 other;

Query Match 0.9%; Score 21.8; DB 18; Length 50;
Best Local Similarity 67.4%; Pred. No. 6.5e+04;
Matches 29; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

Qy 1501 GCATCGGCACCACTTCCAATACGACAGCCGAACTCTGGGCCAT 1543
:|| | || ||| || ||| | ||||||| ||| |||||
Db 43 SCAMCAGCTGCACCTCGGCCACGTCGGCCGAACCCTGAGCCAT 1

RESULT 71
AAT61565/c
ID AAT61565 standard; DNA; 50 BP.
XX
AC AAT61565;
XX
DT 28-OCT-1997 (first entry)
XX
DE scFv antibody library VH back-primer VH 3,5.
XX
KW Human; monoclonal antitumour antibody; peripheral blood lymphocyte;
KW cancer; tumourigenesis; anticancer vaccine; PCR;
KW polymerase chain reaction; ss.
XX
OS Synthetic.
XX
PN WO9702479-A2.
XX
PD 23-JAN-1997.
XX
PF 28-JUN-1996; 96WO-IB01032.
XX
PR 30-JUN-1995; 95US-0497647.
XX
PA (UYYA) UNIV YALE.
XX
PI Cai X, Garen A;
XX
DR WPI; 1997-109061/10.
XX
PT Prodn. of human monoclonal anti-tumour antibodies - by screening a
PT fusion phage library produced using peripheral blood lymphocytes
PT from a cancer patient
XX
PS Example 1; Page 14; 82pp; English.
XX
CC A process for isolating and synthesising human monoclonal anti-tumour
CC antibodies has been produced. The process involves: (a) constructing at
CC least one fusion phage library from the peripheral blood lymphocytes
CC (PBLs) of a cancer patient; (b) screening for anti-tumour antibodies in
CC the phage library in a binding assay with cultured tumour cells of the
CC same type as the patient's tumour; (c) removing extraneous antibodies by
CC absorption against normal human cells; (d) cloning the phage selected in
CC step (b) and (c); (e) assaying the specificity of the cloned phage by
CC incubating the phage with at least two types of cultured normal cells;
CC and (f) further testing the specificity of cloned phage that do not bind
CC to either cell line of cultured normal cells in further binding assays
CC to cultured tumour cells derived from more than one other tumour that is
CC not the patient's tumour. The present sequence represents a VH back-
CC primer involved in the construction of scFv libraries. The
CC antibodies produced can be used for diagnostic and therapeutic
CC applications and for isolating tumour antigens for studying
CC tumourigenesis or for use as anti-cancer vaccines. The human antibodies
CC have low immunogenicity in humans compared to murine monoclonal
CC antibodies (MAbs). Since the antibodies are isolated from fusion phage
CC libraries, their affinity and specificity for a tumour cell line can be
CC improved by genetic manipulations.

```
XX
SQ   Sequence 50 BP; 6 A; 11 C; 20 G; 10 T; 3 other;

  Query Match             0.9%;  Score 21.8;  DB 18;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 6.5e+04;
  Matches   29;  Conservative    1;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    1501 GCATCGGCACCACTTCCAATACGACAGCCGAACTCTGGGCCAT 1543
           :|| | ||  ||| ||    ||| | ||||||| ||| |||||
Db      43 SCAMCAGCTGCACCTCGGCCACGTCGGCCGAACCCTGAGCCAT 1


RESULT 72
AAL28954
ID   AAL28954 standard; DNA; 51 BP.
XX
AC   AAL28954;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #2162.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 2001; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 11 A; 2 C; 5 G; 33 T; 0 other;

  Query Match             0.9%;  Score 21.8;  DB 22;  Length 51;
  Best Local Similarity   70.7%;  Pred. No. 6.5e+04;
  Matches   29;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTT 2148
           | ||||||||| | |  ||| |  ||||| | ||| | ||||
Db      10 TTTTTTTTTTTTGTTTATAGAGTTTTTTAACATAAAGTTTT 50


RESULT 73
AAI76866
ID   AAI76866 standard; DNA; 51 BP.
XX
AC   AAI76866;
XX
DT   09-NOV-2001  (first entry)
XX
DE   Human silent SNP containing nucleic acid SEQ:3807.
XX
KW   Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW   protein therapy; vaccine; probe; diagnostic assay; detection;
KW   quantitation; restorative therapy; polymorphic; ds.
XX
OS   Homo sapiens.
XX
PN   WO200140521-A2.
XX
PD   07-JUN-2001.
XX
PF   30-NOV-2000; 2000WO-US32758.
XX
PR   30-NOV-1999;   99US-0168138.
PR   29-NOV-2000; 2000US-0726173.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-356160/37.
XX
PT   Polymorphic nucleic acid sequences, useful in genetic testing and
PT   therapy -
XX
PS   Claim 1; Page 1216; 2653pp; English.
XX
CC   AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC   sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC   AAM53114 to AAM53329 represent peptides related to human polymorphic
CC   polynucleotide sequences. The sequences can be used in gene and protein
CC   therapy, and in vaccine production. (I) and the polypeptides encoded by
CC   them may be used in the prevention, diagnosis and treatment of diseases
CC   associated with inappropriate expression of polymorphic polypeptides.
CC   For example, (I) may be used to treat disorders by rectifying mutations
CC   or deletions in a patient's genome that affect the activity of
CC   polypeptides by expressing inactive proteins or to supplement the
CC   patients own production of polypeptide. Additionally, (I) and its
CC   complementary sequences may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar nucleic acids
CC   in samples, and therefore which patients may be in need of restorative
CC   therapy. The polypeptides encoded by (I) may be used as antigens in the
CC   production of antibodies specific for polymorphic polypeptides. The
CC   antibodies may also be used to down regulate expression and activity.
CC   The antibodies may also be used as diagnostic agents for detecting the
CC   presence of polymorphic polypeptides in samples.
XX
SQ   Sequence 51 BP; 33 A; 9 C; 5 G; 4 T; 0 other;

  Query Match             0.9%;  Score 21.8;  DB 22;  Length 51;
  Best Local Similarity   70.7%;  Pred. No. 6.5e+04;
  Matches   29;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy     918 TAGAAAGAAAACAATTGACAAAATCCTCAAATGTAAACTCA 958
           |||||| || | ||   | |||| ||||||| | |||  ||
```

```
Db     10 TAGAAAAAACAAAAAACAAAAAACCCTCAAAGGAAAAAACA 50


RESULT 74
AAV76001
ID   AAV76001 standard; DNA; 55 BP.
XX
AC   AAV76001;
XX
DT   16-MAR-1999  (first entry)
XX
DE   Staphylococcus aureus contig SEQ ID #1690.
XX
KW   Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW   cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW   skin infection; surgical wound infection; scalded skin syndrome;
KW   toxic shock syndrome; ds.
XX
OS   Staphylococcus aureus.
XX
PN   EP786519-A2.
XX
PD   30-JUL-1997.
XX
PF   07-JAN-1997;   97EP-0100117.
XX
PR   05-JAN-1996;   96US-0009861.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Barash SC,  Choi GH,  Dillon PJ,  Fannon MR,  Kunsch CA;
PI   Rosen CA;
XX
DR   WPI; 1997-374922/35.
XX
PT   Polynucleotide(s) and proteins derived from Staphylococcus aureus  -
PT   stored on computer readable medium and used in the production of
PT   anti-S.aureus vaccines
XX
PS   Claim 1; Page 2034; 3271pp; English.
XX
CC   This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC   of the invention. The DNA sequences are recorded on a computer readable
CC   medium, preferably selected from a floppy or hard disk, random access
CC   memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC   the S.aureus DNA sequences allows putative functions to be assigned so
CC   that protein-encoding or regulatory regions of commercial, therapeutic or
CC   industrial importance can be obtained. Specifically, sequences which are
CC   likely to encode antigens have been identified and these polypeptides can
CC   be used in a vaccine composition against S.aureus infection. The
CC   polypeptides can also be used in a kit for the immunodetection of
CC   S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC   including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC   skin and surgical wound infections, scalded skin syndrome, toxic shock
CC   syndrome, etc. Organisms transformed with the DNA sequences can be used
CC   for recombinant production of the polypeptides. The new DNA sequences
CC   (and their fragments) are useful as primers or probes for isolating
CC   homologues of any of the S.aureus DNA sequences contained on the
CC   computer readable medium.
XX
SQ   Sequence 55 BP; 24 A; 9 C; 6 G; 16 T; 0 other;

  Query Match             0.9%;  Score 21.8;  DB 18;  Length 55;
  Best Local Similarity   70.7%;  Pred. No. 6.8e+04;
  Matches   29;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   2174 CCAAACTTTAAAATGCGAAATTATTGGTTGGTGTGAAGAAA 2214
           |||||||||| |||  |||||  |||||   |   |||| |
Db      1 CCAAACTTTATAATATGAAATGCTTGGTAATTACAAAGATA 41


RESULT 75
AAF56143/c
ID   AAF56143 standard; DNA; 60 BP.
XX
AC   AAF56143;
XX
DT   17-APR-2001  (first entry)
XX
DE   Staphylococcus aureus agr enhancer region #1.
XX
KW   Staphylococcus aureus; SarA; staphylococcal accessory regulator A;
KW   agr; accessory gene regulator; antibacterial; SarA inhibitor;
KW   virulence gene; staphylococcal infection; ds.
XX
OS   Staphylococcus aureus.
XX
PN   WO200103686-A2.
XX
PD   18-JAN-2001.
XX
PF   07-JUL-2000; 2000WO-US18525.
XX
PR   08-JUL-1999;   99US-0142793.
XX
PA   (UYAR-) UNIV ARKANSAS.
XX
PI   Hurlburt BK,  Smeltzer MS,  Rechtin TM;
XX
DR   WPI; 2001-112567/12.
XX
PT   Identifying inhibitors of staphylococcal SarA (accessory regulator)
PT   which are useful for treating staphylococcal infections, comprises
PT   using specific binding sites of SarA protein on an accessory gene
PT   regulator locus  -
XX
PS   Example; Fig 2; 79pp; English.
XX
CC   The present sequence is given in a specification relating to a method for
CC   identifying inhibitors of SarA (staphylococcal accessory regulator)
CC   function involved in the expression of Staphylococcal virulence genes.
CC   The method comprises contacting a candidate inhibitor with a SarA
CC   binding site of the agr (accessory gene regulator) locus in solution
CC   and assessing the binding of the candidate inhibitor to the SarA
CC   binding site of the agr locus. The identified inhibitors are useful for
CC   preventing and treating staphylococcal infections.
XX
SQ   Sequence 60 BP; 17 A; 5 C; 4 G; 34 T; 0 other;

  Query Match             0.9%;  Score 21.8;  DB 22;  Length 60;
  Best Local Similarity   61.4%;  Pred. No. 7.1e+04;
  Matches   35;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy   1709 AATGTCATTACATAGAACACTTCGGACACAGGAAAAATAAACGTGGATTTTAAAAAA 1765
           ||| |  | ||||| || | ||   |   | ||| || |||||   | || | ||||
Db     60 AATATTTTAACATAAAAAAATTTACAGTTAAGAATAAAAAACGACTAGTTAAGAAAA 4


RESULT 76
ABN35167
ID   ABN35167 standard; DNA; 60 BP.
XX
AC   ABN35167;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:7915.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
```

XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 7915; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 24 A; 10 C; 18 G; 8 T; 0 other;

Query Match 0.9%; Score 21.8; DB 24; Length 60;
Best Local Similarity 70.7%; Pred. No. 7.1e+04;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 974 TCACACAAGAAGCCAGAGATCTGCTTAAAAAGCTGCTGAAA 1014
| ||| | ||||| ||||| | ||| |||| ||||||
Db 5 TAACAGAGGAAGCGAGAGAAGGTATCAAACAGCTACTGAAA 45

RESULT 77
ABN43764/c
ID ABN43764 standard; DNA; 60 BP.
XX
AC ABN43764;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:16512.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.

XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 16512; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 15 A; 13 C; 15 G; 17 T; 0 other;

Query Match 0.9%; Score 21.8; DB 24; Length 60;
Best Local Similarity 61.4%; Pred. No. 7.1e+04;
Matches 35; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 215 AACTTGGCATGGAACATTGTGAGAAATTTGAAATCTCAGAAACTAGTGTGAACAGAG 271
|| ||||| || ||| || | || ||||||| | | ||| || || ||
Db 59 AATTTGGCCAGGCACAGTGGTTCGCACCTGTAATCTCAACATATTGTGAGACCAAAG 3

RESULT 78
ABN50818
ID ABN50818 standard; DNA; 60 BP.
XX
AC ABN50818;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:23566.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.

```
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes -
XX
PS   Example 1; SEQ ID 23566; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 11 A; 12 C; 18 G; 19 T; 0 other;

  Query Match             0.9%;  Score 21.8;  DB 24;  Length 60;
  Best Local Similarity   65.3%;  Pred. No. 7.1e+04;
  Matches   32;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   1015 AGAAATGCTGCTTCTCGTCTGGGAGCTGGTCCTGGGGACGCTGGAGAAG 1063
           ||||||| |||||||  | | |  ||   |||  |  ||| ||| | ||
Db      3 AGAAATGTTGCTTCTACTTTTGAGGCGCTTCCCTGAAACGGTGGTGCAG 51


RESULT 79
AAL30101/c
ID   AAL30101 standard; DNA; 51 BP.
XX
AC   AAL30101;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #3309.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 2335; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 10 A; 1 C; 4 G; 36 T; 0 other;

  Query Match             0.9%;  Score 21.6;  DB 22;  Length 51;
  Best Local Similarity   68.2%;  Pred. No. 7.3e+04;
  Matches   30;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1757 TTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAA 1800
           || |||||||||  || |     |||||||||   |||  ||||
Db     45 TTAAAAAAATCATACATTCACAAAAAAAAAAAAAAAAAAAAAA 2


RESULT 80
AAL30786/c
ID   AAL30786 standard; DNA; 51 BP.
XX
AC   AAL30786;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #3994.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
```

PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 2534; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 51 BP; 8 A; 2 C; 3 G; 38 T; 0 other;

Query Match 0.9%; Score 21.6; DB 22; Length 51;
Best Local Similarity 68.2%; Pred. No. 7.3e+04;
Matches 30; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1759 TAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
| ||| || || |||| | ||||||||| ||| |||| |
Db 46 TTAAACATAAAACAATTGCTTAAAAAAAAAAAAAAAAAAAAAAA 3

RESULT 81
AAI76504/c
ID AAI76504 standard; DNA; 51 BP.
XX
AC AAI76504;
XX
DT 09-NOV-2001 (first entry)
XX
DE Human silent SNP containing nucleic acid SEQ:3445.
XX
KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW protein therapy; vaccine; probe; diagnostic assay; detection;
KW quantitation; restorative therapy; polymorphic; ds.
XX
OS Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.
XX
PF 30-NOV-2000; 2000WO-US32758.
XX
PR 30-NOV-1999; 99US-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-356160/37.
XX
PT Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
PS Claim 1; Page 1105; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.
XX
SQ Sequence 51 BP; 13 A; 10 C; 5 G; 23 T; 0 other;

Query Match 0.9%; Score 21.6; DB 22; Length 51;
Best Local Similarity 68.2%; Pred. No. 7.3e+04;
Matches 30; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1278 TGTACTTGAAAGTGTGAAAGAAAAGTTTTCCTTTGAACCAAAAA 1321
||||| ||| ||||||| ||| ||| || || |||||
Db 45 TGTACACTAAAATGTGAAATTGAAGGTTTTGTTAAAAAGAAAAA 2

RESULT 82
AAI76505/c
ID AAI76505 standard; DNA; 51 BP.
XX
AC AAI76505;
XX
DT 09-NOV-2001 (first entry)
XX
DE Human silent SNP containing nucleic acid SEQ:3446.
XX
KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW protein therapy; vaccine; probe; diagnostic assay; detection;
KW quantitation; restorative therapy; polymorphic; ds.
XX
OS Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.
XX
PF 30-NOV-2000; 2000WO-US32758.
XX
PR 30-NOV-1999; 99US-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-356160/37.

```
XX
PT   Polymorphic nucleic acid sequences, useful in genetic testing and
PT   therapy -
XX
PS   Claim 1; Page 1105; 2653pp; English.
XX
CC   AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC   sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC   AAM53114 to AAM53329 represent peptides related to human polymorphic
CC   polynucleotide sequences. The sequences can be used in gene and protein
CC   therapy, and in vaccine production. (I) and the polypeptides encoded by
CC   them may be used in the prevention, diagnosis and treatment of diseases
CC   associated with inappropriate expression of polymorphic polypeptides.
CC   For example, (I) may be used to treat disorders by rectifying mutations
CC   or deletions in a patient's genome that affect the activity of
CC   polypeptides by expressing inactive proteins or to supplement the
CC   patients own production of polypeptide. Additionally, (I) and its
CC   complementary sequences may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar nucleic acids
CC   in samples, and therefore which patients may be in need of restorative
CC   therapy. The polypeptides encoded by (I) may be used as antigens in the
CC   production of antibodies specific for polymorphic polypeptides. The
CC   antibodies may also be used to down regulate expression and activity.
CC   The antibodies may also be used as diagnostic agents for detecting the
CC   presence of polymorphic polypeptides in samples.
XX
SQ   Sequence 51 BP; 12 A; 10 C; 6 G; 23 T; 0 other;

  Query Match             0.9%;  Score 21.6;  DB 22;  Length 51;
  Best Local Similarity   68.2%;  Pred. No. 7.3e+04;
  Matches   30;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1278 TGTACTTGAAAGTGTGAAAGAAAAGTTTTCCTTTGAACCAAAAA 1321
            |||||   ||| |||||||   ||| |||  ||  ||  |||||
Db     45 TGTACACTAAAATGTGAAACTGAAGGTTTTGTTAAAAAGAAAAA 2


RESULT 83
AAT23964/c
ID   AAT23964 standard; cDNA to mRNA; 54 BP.
XX
AC   AAT23964;
XX
DT   27-AUG-1996  (first entry)
XX
DE   Human gene signature HUMGS05918.
XX
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
XX
OS   Homo sapiens.
XX
PN   WO9514772-A1.
XX
PD   01-JUN-1995.
XX
PF   11-NOV-1994;   94WO-JP01916.
XX
PR   12-NOV-1993;   93JP-0355504.
XX
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
XX
PI   Matsubara K,  Okubo K;
XX
DR   WPI; 1995-206931/27.
XX
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
XX
PS   Claim 1; Page 1497; 2245pp; Japanese.
XX
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in AAT19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
XX
SQ   Sequence 54 BP; 19 A; 6 C; 8 G; 21 T; 0 other;

  Query Match             0.9%;  Score 21.6;  DB 16;  Length 54;
  Best Local Similarity   63.5%;  Pred. No. 7.5e+04;
  Matches   33;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy   2086 TTTTACGTGCAAACAACCTGAATCTTTTTTTTATATAAATATATATTTTTCA 2137
           ||| | |  | | |||  | || | ||| ||| ||||| | ||||  | | |
Db     54 TTTGAAGAACGACCAAAATCAAACCTTTATTTTTATAACTGTATAAATGTGA 3


RESULT 84
ABN36856
ID   ABN36856 standard; DNA; 60 BP.
XX
AC   ABN36856;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:9604.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes -
XX
PS   Example 1; SEQ ID 9604; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
```

```
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 22 A; 12 C; 18 G; 8 T; 0 other;

  Query Match             0.9%;  Score 21.6;  DB 24;  Length 60;
  Best Local Similarity   63.5%;  Pred. No. 8e+04;
  Matches   33;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy     245 AAATCTCAGAAACTAGTGTGAACAGAGGGCCAGAAAAAATCAGACCAGAATG 296
           ||| |  || |||   ||  | |  |||||  ||||  |||||| | || ||
Db       1 AAAGCAGAGGAACCCATGACATCCAAGGGCGTGAAAGGATCAGAGCTGACTG 52


RESULT 85
AAV69937
ID   AAV69937 standard; DNA; 48 BP.
XX
AC   AAV69937;
XX
DT   11-FEB-1999  (first entry)
XX
DE   Chlorella virus promoter consensus sequence cvp-6.
XX
KW   Promoter; Chlorella virus; structural gene; lac operon;
KW   protein expression; ss.
XX
OS   Chlorella sp.
XX
PN   WO9842822-A1.
XX
PD   01-OCT-1998.
XX
PF   21-MAR-1998;   98WO-US05655.
XX
PR   21-MAR-1997;   97US-0821559.
XX
PA   (BION-) BIONEBRASKA INC.
XX
PI   Xia Y;
XX
DR   WPI; 1998-609893/51.
XX
PT   Promoter sequences derived from Chlorella virus - are useful for
PT   heterologous protein production in plants
XX
PS   Example 6; Page 15; 46pp; English.
XX
CC   AAV69935-46 represent promoter consensus sequences. The specification
CC   describes promoter sequences derived from Chlorella virus. These
CC   promoter sequences direct transcription of a structural gene. A lac
CC   operon can be linked operatively to the promoters. The promoters can
CC   be used in constructs for promoting high level expression of proteins
CC   e.g. heterologous proteins, in a host, especially plants such tobacco,
CC   wheat.
XX
SQ   Sequence 48 BP; 17 A; 7 C; 9 G; 15 T; 0 other;

  Query Match             0.9%;  Score 21.4;  DB 19;  Length 48;
  Best Local Similarity   66.0%;  Pred. No. 8e+04;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy     2233 CTTCTCTTGGTGAAATAATAAAATGCAAATGAATCATTGTTAACACA 2279
            ||| || || ||| || |||||   |||||||  |  |||    |||
Db        2 CTTATCGTGTTGATATGATAAATGACAAATGATACGCTGTATCAACA 48


RESULT 86
AAS21106
ID   AAS21106 standard; DNA; 48 BP.
XX
AC   AAS21106;
XX
DT   20-MAR-2002  (first entry)
XX
DE   (GGA)16 DNA purification oligonucleotide.
XX
KW   ss; DNA purification; triple helix; plasmid purification.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   repeat_region   1..48
FT                   /*tag=  a
FT                   /rpt_type= "TANDEM"
FT   repeat_unit     1..3
FT                   /*tag=  b
FT                   /note= "GAA repeat type"
XX
PN   WO200192511-A2.
XX
PD   06-DEC-2001.
XX
PF   25-MAY-2001; 2001WO-US17122.
XX
PR   26-MAY-2000; 2000US-0580923.
XX
PA   (AVET ) AVENTIS PHARMA SA.
XX
PI   Crouzet J,  Scherman D,  Wils P,  Blanche F,  Cameron B;
XX
DR   WPI; 2002-097772/13.
XX
PT   Purifying double-stranded (ds) DNA from a solution containing dsDNA and
PT   other components, comprises passing the solution through a support
PT   comprising a covalently coupled oligonucleotide able to form a triple
PT   helix with the dsDNA  -
XX
PS   Example 7; Page 20; 40pp; English.
XX
CC   This invention comprises a method of purifying double-stranded DNA from
CC   a solution containing the double-stranded DNA mixed with other
CC   components, comprising passing the solution through a support comprising
CC   a covalently coupled oligonucleotide capable of forming a triple helix
CC   with the double-stranded DNA by hybridisation with a specific sequence
CC   present in the double-stranded DNA. The method is useful for purifying
CC   double-stranded DNA contained in a solution and mixed with other
CC   components. The new method is a simple, rapid and effective method for
CC   DNA purification, and makes it possible to obtain especially high
CC   purities with high yields. The method enables DNA to be purified from
CC   complex mixtures comprising other nucleic acids, proteins, endotoxins,
CC   nucleases and the like. The supports may be readily recycled, and the
CC   DNAs obtained display improved properties to pharmaceutical safety.
CC   Further, the method entails only one step contrary to prior art.
CC   The present sequence represents a DNA sequence contained within the
CC   plasmid pXL2725. This sequence is used for purification of this
CC   plasmid using the method of the invention.
XX
SQ   Sequence 48 BP; 16 A; 0 C; 32 G; 0 U; 0 other;

  Query Match             0.9%;  Score 21.4;  DB 24;  Length 48;
```

```
  Best Local Similarity   66.0%;  Pred. No. 8e+04;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    123 GGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGG 169
          |||  ||   ||||| | ||     ||||| |||  ||||||| ||
Db      1 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 47


RESULT 87
AAD24852
ID   AAD24852 standard; DNA; 48 BP.
XX
AC   AAD24852;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Chlorella virus cvp-6 promoter DNA consensus.
XX
KW   Gene expression; heterologous enzyme; hormone; structural protein;
KW   Escherichia coli cell; tobacco; wheat; ds.
XX
OS   Chlorella virus.
XX
PN   US6316224-B1.
XX
PD   13-NOV-2001.
XX
PF   21-SEP-1999;   99US-0400541.
XX
PR   21-MAR-1997;   97US-0821559.
PR   21-MAR-1998;   98WO-US05655.
XX
PA   (BION-) BIONEBRASKA INC.
XX
PI   Xia Y;
XX
DR   WPI; 2002-054589/07.
XX
PT   Chlorella virus promoters, useful for controlling expression of
PT   hormones, enzymes and structural proteins in plant and Escherichia coli
PT   cells  -
XX
PS   Example 6; Column 9-10; 24pp; English.
XX
CC   The invention relates to isolated, inducible promoter sequences derived
CC   from Chlorella viruses. The invention also relates to gene constructs
CC   comprising a promoter sequence of the invention operably linked to a DNA
CC   sequence of a structural gene. The promoter is used for controlling gene
CC   expression in host cells. In particular, it is used to control expression
CC   of genes encoding heterologous enzymes, hormones, structural proteins,
CC   glucagon-like peptide 1, growth hormone releasing factor, parathyroid
CC   hormone, carbonic anhydrase, beta-galactosidase, chloramphenicol acetyl
CC   transferase or glutathione acetyltransferase in plants cells and
CC   Escherichia coli cells. The promoters can also promote high levels of
CC   expression in plants especially tobacco and wheat. The present sequence
CC   is Chlorella virus cvp-6 promoter DNA consensus.
XX
SQ   Sequence 48 BP; 17 A; 7 C; 9 G; 15 T; 0 other;

  Query Match             0.9%;  Score 21.4;  DB 24;  Length 48;
  Best Local Similarity   66.0%;  Pred. No. 8e+04;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   2233 CTTCTCTTGGTGAAATAATAAAATGCAAATGAATCATTGTTAACACA 2279
          ||| || || ||| || |||||   |||||||  |  |||    |||
Db      2 CTTATCGTGTTGATATGATAAATGACAAATGATACGCTGTATCAACA 48


RESULT 88
AAL31468
ID   AAL31468 standard; DNA; 51 BP.
XX
AC   AAL31468;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #4676.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 2731; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 13 A; 6 C; 6 G; 26 T; 0 other;

  Query Match             0.9%;  Score 21.4;  DB 22;  Length 51;
  Best Local Similarity   80.6%;  Pred. No. 8.3e+04;
  Matches   25;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy   2100 AACCTGAATCTTTTTTTTATATAAATATATA 2130
          |||| |||||||||||||| | | ||| | ||
Db      4 AACCAGAATCTTTTTTTTTTTTTAATCTGTA 34


RESULT 89
AAI78865/c
ID   AAI78865 standard; DNA; 51 BP.
XX
AC   AAI78865;
XX
DT   09-NOV-2001  (first entry)
XX
```

DE Human silent SNP containing nucleic acid SEQ:5806.
XX
KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW protein therapy; vaccine; probe; diagnostic assay; detection;
KW quantitation; restorative therapy; polymorphic; ds.
XX
OS Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.
XX
PF 30-NOV-2000; 2000WO-US32758.
XX
PR 30-NOV-1999; 99US-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-356160/37.
XX
PT Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
PS Claim 1; Page 2287; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.
XX
SQ Sequence 51 BP; 18 A; 2 C; 5 G; 26 T; 0 other;

Query Match 0.9%; Score 21.4; DB 22; Length 51;
Best Local Similarity 66.0%; Pred. No. 8.3e+04;
Matches 31; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 2097 AACAACCTGAATCTTTTTTTTATATAAATATATATTTTTCAAATAGA 2143
||||| | || | | || | |||| |||||||| | | ||| |
Db 48 AACAAAATATATATATATTCTCTATATATATATATATAGCTGATAAA 2

RESULT 90
AAH93331/c
ID AAH93331 standard; DNA; 53 BP.
XX
AC AAH93331;
XX
DT 04-OCT-2001 (first entry)
XX
DE Plasmodium falciparum MAL3PC polynucleotide SEQ ID NO 52.
XX
KW Human; antisense-therapy; gene-therapy; diagnostic; forensic;
KW gene mapping; ds.
XX
OS Plasmodium falciparum.
XX
PN WO200152616-A2.
XX
PD 26-JUL-2001.
XX
PF 22-DEC-2000; 2000WO-US35190.
XX
PR 23-DEC-1999; 99US-0471275.
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-451890/48.
XX
PT Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use -
XX
PS Example 4; Page 120; 135pp; English.
XX
CC The invention relates to an isolated human polynucleotide (AAH75398)
CC encoding a novel polypeptide (AAG64527) useful in antisense-therapy and
CC gene-therapy, in diagnostics, forensics, gene mapping and identification
CC of mutations responsible for genetic disorders and other traits.
CC Polynucleotide sequences with potential homology were also identified
CC (AAH93283-AAH93356).
XX
SQ Sequence 53 BP; 27 A; 2 C; 0 G; 24 T; 0 other;

Query Match 0.9%; Score 21.4; DB 22; Length 53;
Best Local Similarity 71.8%; Pred. No. 8.4e+04;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2111 TTTTTTTATATAAATATATATTTTTCAAATAGATTTTTG 2149
| | | |||||| |||||||| | | |||| || | ||
Db 43 TATATATATATATATATATATATATATAATATATATGTG 5

RESULT 91
AAC11350
ID AAC11350 standard; cDNA; 59 BP.
XX
AC AAC11350;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 15425.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX

PS Claim 1; SEQ ID 15425; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 59 BP; 22 A; 8 C; 8 G; 21 T; 0 other;

Query Match 0.9%; Score 21.4; DB 21; Length 59;
Best Local Similarity 66.0%; Pred. No. 8.9e+04;
Matches 31; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 2090 ACGTGCAAACAACCTGAATCTTTTTTTTATATAAATATATATTTTTC 2136
|| | || | | | ||||||| |||||||||||| | ||
Db 10 ACATTGAATGCAAACGTGTATTTTTTTTAATATAAATATATAACTGTC 56

RESULT 92
ABN32960/c
ID ABN32960 standard; DNA; 60 BP.
XX
AC ABN32960;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:5708.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 5708; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 16 A; 11 C; 13 G; 20 T; 0 other;

Query Match 0.9%; Score 21.4; DB 24; Length 60;
Best Local Similarity 66.0%; Pred. No. 9e+04;
Matches 31; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 2174 CCAAACTTTAAAATGCGAAATTATTGGTTGGTGTGAAGAAAGCCAGA 2220
|||| ||| | | | | || | ||| || ||||||||||||
Db 59 CCAAGTTTTCATTTACTATGCCATCGCTTGCTGCAAAGAAAGCCAGA 13

RESULT 93
ABN37046
ID ABN37046 standard; DNA; 60 BP.
XX
AC ABN37046;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:9794.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 9794; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a

CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 16 A; 12 C; 14 G; 18 T; 0 other;

Query Match 0.9%; Score 21.4; DB 24; Length 60;
Best Local Similarity 71.8%; Pred. No. 9e+04;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1256 GTTTTACATATGTGGCTCCATCTGTACTTGAAAGTGTGA 1294
|| | |||| ||| | || || ||||| |||||||
Db 9 GTGTAACATTCAGGGCACTATGTGAACTTGTAAGTGTGA 47

RESULT 94
ABN37096/c
ID ABN37096 standard; DNA; 60 BP.
XX
AC ABN37096;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:9844.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 9844; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a

CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 15 A; 17 C; 13 G; 15 T; 0 other;

Query Match 0.9%; Score 21.4; DB 24; Length 60;
Best Local Similarity 66.0%; Pred. No. 9e+04;
Matches 31; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1011 GAAAAGAAATGCTGCTTCTCGTCTGGGAGCTGGTCCTGGGGACGCTG 1057
|||||| |||||| | | |||| |||| | ||| ||| ||
Db 47 GAAAAGGAATGCTCATCTTGCCCTGGATGCTGCTGATGGCGACAATG 1

RESULT 95
ABN50863/c
ID ABN50863 standard; DNA; 60 BP.
XX
AC ABN50863;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:23611.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 23611; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a

```
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 20 A; 15 C; 11 G; 14 T; 0 other;

  Query Match             0.9%;  Score 21.4;  DB 24;  Length 60;
  Best Local Similarity   71.8%;  Pred. No. 9e+04;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   2191 AAATTATTGGTTGGTGTGAAGAAAGCCAGACAACTTCTG 2229
          |||| || |||   || ||||| |||| || | | ||||
Db     39 AAATGATGGGTCTATGAGAAGATAGCCTGATAGCCTCTG 1


RESULT 96
AAH44690/c
ID   AAH44690 standard; DNA; 41 BP.
XX
AC   AAH44690;
XX
DT   07-DEC-2001  (first entry)
XX
DE   Human type-I aminoacyl tRNA synthetase 10 probe 1 SEQ ID NO:8.
XX
KW   Human; type-I aminoacyl tRNA synthetase 10; malignant tumour;
KW   haemopathy; human immunodeficiency virus; HIV infection;
KW   immunological disease; inflammation; probe; ss.
XX
OS   Homo sapiens.
XX
PN   CN1301715-A.
XX
PD   04-JUL-2001.
XX
PF   27-DEC-1999;   99CN-0125371.
XX
PR   27-DEC-1999;   99CN-0125371.
XX
PA   (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2001-550469/62.
XX
PT   New polypeptide-I-type aminoacyl/tRNA synthetase 10 and encoding
PT   polynucleotide useful for treating tumor, hemopathy, infection and
PT   immunological disease -
XX
PS   Example 7; Page 20 (Disclosure); 32pp; Chinese.
XX
CC   The present invention describes the human type-I aminoacyl tRNA
CC   synthetase 10 protein. Also described are polynucleotides encoding the
CC   type-I amino acyl tRNA synthetase 10 protein, and a DNA recombination
CC   process to produce the protein. The protein can be used for treating
CC   various diseases, such as malignant tumour, haemopathy, human
CC   immunodeficiency virus infection, immunological diseases and various
CC   inflammations. The present sequence represents a probe for type-I
CC   aminoacyl tRNA synthetase 10, which is used in an example from the
CC   present invention.
XX
SQ   Sequence 41 BP; 3 A; 5 C; 3 G; 30 T; 0 other;

  Query Match             0.9%;  Score 21.2;  DB 22;  Length 41;
  Best Local Similarity   76.5%;  Pred. No. 8.4e+04;
  Matches   26;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   2045 AAAATGGGAAGGCAAAGACAAAAGAAACTTACCA 2078
          ||||  | || | ||||| ||||| ||||||| |
Db     37 AAAAAAGAAAAGAAAAGAAAAAAGGAACTTACAA 4


RESULT 97
AAF29312/c
ID   AAF29312 standard; DNA; 48 BP.
XX
AC   AAF29312;
XX
DT   18-APR-2001  (first entry)
XX
DE   Primer base sequence used to illustrate primer selection method.
XX
KW   Primer; optimum sequence; differential display; ss.
XX
OS   Synthetic.
XX
PN   JP2000308487-A.
XX
PD   07-NOV-2000.
XX
PF   30-MAR-1999;   99JP-0088410.
XX
PR   30-MAR-1999;   99JP-0088410.
XX
PA   (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
DR   WPI; 2001-046077/06.
XX
PT   Selection of primer base for optimizing primer selection comprises
PT   obtaining an optimum sequence for differential display from an
PT   expression gene data base -
XX
PS   Disclosure; Fig 9; 13pp; Japanese.
XX
CC   This invention relates to a method for selecting the sequence of a
CC   primer. The method comprises obtaining an optimum sequence for
CC   differential display from an expression gene data base, and using the
CC   base sequences most frequently expressed as the primer candidates in the
CC   order of frequency. The optimum primer group characterised by the use of
CC   genetic algorithm from the primer candidates is selected. The method is
CC   used for selecting a primer sequence quickly. The present sequence
CC   represents a primer used in an illustration of the method of the
CC   invention.
XX
SQ   Sequence 48 BP; 18 A; 2 C; 5 G; 23 T; 0 other;

  Query Match             0.9%;  Score 21.2;  DB 22;  Length 48;
  Best Local Similarity   69.0%;  Pred. No. 9e+04;
  Matches   29;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   2097 AACAACCTGAATCTTTTTTTTATATAAATATATATTTTTCAA 2138
          || ||  |||||||||   || || ||| | |  || |||||
Db     47 AAGAATTTGAATCTTTCAATTTTAAAAAAAAAATTTCTTCAA 6


RESULT 98
AAI77361/c
ID   AAI77361 standard; DNA; 50 BP.
XX
AC   AAI77361;
XX
DT   09-NOV-2001  (first entry)
XX
```

DE Human silent SNP containing nucleic acid SEQ:4302.
XX
KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW protein therapy; vaccine; probe; diagnostic assay; detection;
KW quantitation; restorative therapy; polymorphic; ds.
XX
OS Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.
XX
PF 30-NOV-2000; 2000WO-US32758.
XX
PR 30-NOV-1999; 99US-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-356160/37.
XX
PT Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
PS Claim 1; Page 1827; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.
XX
SQ Sequence 50 BP; 12 A; 13 C; 7 G; 18 T; 0 other;

Query Match 0.9%; Score 21.2; DB 22; Length 50;
Best Local Similarity 69.0%; Pred. No. 9.2e+04;
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy    895 GCACCCCCATTCACTGGGGAGAATAGAAAGAAAACAATTGAC 936
          ||| || ||| ||||||| ||| | ||||  || |  || |
Db     46 GCATACCAATTTACTGGGGGGAAAAAAAAGTTAAGAGATGCC 5
```

RESULT 99
AAA77364/c
ID AAA77364 standard; cDNA; 51 BP.
XX
AC AAA77364;
XX
DT 16-NOV-2000 (first entry)
XX
DE Human clone cg44911913 polymorphic site, SEQ ID NO:1047.
XX
KW Human; single nucleotide polymorphism; SNP; chromosome 15;
KW detection; identification; gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace (26,G)
FT /*tag= a
XX
PN WO200029623-A2.
XX
PD 25-MAY-2000.
XX
PF 17-NOV-1999; 99WO-US27293.
XX
PR 17-NOV-1998; 98US-0109024.
PR 16-NOV-1999; 99US-0109024.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach MD;
XX
DR WPI; 2000-387826/33.
XX
PT Human nucleic acids containing single nucleotide polymorphisms, useful
PT for treating a subject suffering, or at risk from a pathology due to
PT the presence of a sequence polymorphism -
XX
PS Claim 1; Page 474; 543pp; English.
XX
CC Sequences AAA76318-A77509 represent 1192 human nucleic acid sequences
CC which contain single nucleotide polymorphisms (SNPs). Sequences 1 to
CC 1112 (AAA76318-A77429) are consecutive pairs of nucleotides which
CC contain silent SNPs. Sequences 1113 to 1192 (AAA77430-A77509) are
CC consecutive pairs of nucleotides containing SNPs which result in changes
CC in the corresponding amino acid sequences (AAB11749-B11828). The SNPs in
CC sequences 1113 to 1128 (AAA77430-A77445) lead to conservative amino acid
CC changes, while those in sequences 1129 to 1186 (AAA77446-A77503) result
CC in non- conservative changes. The SNPs in sequences 1187 to 1192
CC (AAA77504-A77509) generate frameshift mutations. The invention also
CC relates to a method of detecting a polymorphic site in a nucleic acid and
CC a method of determining the relatedness of two nucleic acids. It also
CC encompasses peptides containing polymorphic sites, antibodies raised
CC against such peptides, and a method of detecting polymorphic
CC proteins/peptides using the antibodies. The nucleic acids are useful for
CC gene therapy of an individual having, suspected of having, or at risk of
CC developing a pathological condition due to the presence of a sequence
CC polymorphism. Such treatment would comprise administration of the
CC wild-type nucleic acid sequence. Antibodies raised against polymorphic
CC peptides can also be used in the treatment of such individuals.
XX
SQ Sequence 51 BP; 0 A; 30 C; 12 G; 9 T; 0 other;

Query Match 0.9%; Score 21.2; DB 21; Length 51;
Best Local Similarity 64.0%; Pred. No. 9.3e+04;
Matches 32; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy    114 CATAGACCTGGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGG 163
          ||  || | || ||||     ||||| |||| |  |||| | || || ||
Db     51 CAGGGAGCAGGGCCAGGAGCGGGACGAGGGCGCCCAGGAGGGGCCGGGGG 2
```

RESULT 100
AAL31990/c
ID AAL31990 standard; DNA; 51 BP.
XX
AC AAL31990;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #5198.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;

KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 2882; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 51 BP; 11 A; 4 C; 4 G; 32 T; 0 other;

Query Match 0.9%; Score 21.2; DB 22; Length 51;
Best Local Similarity 64.0%; Pred. No. 9.3e+04;
Matches 32; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy   1749 ACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAA 1798
          | |||| || ||  |||  ||| |    | ||||||||||   |||  ||
Db     50 ATGTGGTTTCTATTAAAAAAATAATACATCCAAAAAAAAAAAAAAAAAAA 1
```

Search completed: April 19, 2003, 09:03:18
Job time : 554 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: April 19, 2003, 06:56:13 ; Search time 96 Seconds
(without alignments)
7494.413 Million cell updates/sec

Title: US-09-920-677-3
Perfect score: 2346
Sequence: 1 gcacgaggctgcggcgggtc.....gtaacccagctgcggagcct 2346

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 635134

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	23.6	1.0	60	4	US-09-461-697-231	Sequence 231, App
	2	23.4	1.0	54	1	US-08-198-094-72	Sequence 72, Appl
	3	23.4	1.0	54	5	PCT-US95-02087-72	Sequence 72, Appl
	4	23.4	1.0	58	4	US-08-860-038-15	Sequence 15, Appl
	5	23.4	1.0	58	4	US-09-580-923-15	Sequence 15, Appl
	6	22.6	1.0	54	2	US-08-823-516-162	Sequence 162, App
	7	22.4	1.0	44	1	US-08-171-389-63	Sequence 63, Appl
	8	22.4	1.0	44	1	US-08-123-936-63	Sequence 63, Appl
	9	22.4	1.0	44	2	US-08-475-228A-63	Sequence 63, Appl
	10	22.4	1.0	44	3	US-08-482-080A-63	Sequence 63, Appl
	11	22.4	1.0	44	4	US-09-354-947-63	Sequence 63, Appl
	12	22.4	1.0	44	5	PCT-US93-12388-63	Sequence 63, Appl
	13	22.2	0.9	45	1	US-08-171-389-212	Sequence 212, App
	14	22.2	0.9	45	1	US-08-123-936-212	Sequence 212, App
	15	22.2	0.9	45	2	US-08-475-228A-212	Sequence 212, App
	16	22.2	0.9	45	3	US-08-482-080A-212	Sequence 212, App
	17	22.2	0.9	45	4	US-09-354-947-212	Sequence 212, App
	18	22.2	0.9	45	5	PCT-US93-12388-212	Sequence 212, App
c	19	22.2	0.9	47	4	US-09-641-638-880	Sequence 880, App
	20	22	0.9	42	3	US-08-975-703-31	Sequence 31, Appl
	21	22	0.9	42	4	US-09-515-884-31	Sequence 31, Appl
c	22	22	0.9	58	4	US-08-860-038-16	Sequence 16, Appl
c	23	22	0.9	58	4	US-09-580-923-16	Sequence 16, Appl
	24	22	0.9	60	4	US-09-021-560-5	Sequence 5, Appli
c	25	21.8	0.9	50	3	US-08-983-607-7	Sequence 7, Appli
c	26	21.8	0.9	50	3	US-08-983-607-44	Sequence 44, Appl
	27	21.4	0.9	48	2	US-08-821-559A-12	Sequence 12, Appl
	28	21.4	0.9	48	4	US-09-400-541-12	Sequence 12, Appl
	29	21.4	0.9	48	4	US-09-580-923-34	Sequence 34, Appl
	30	21.4	0.9	48	4	US-09-899-999-12	Sequence 12, Appl
	31	21.4	0.9	55	1	US-08-113-646A-41	Sequence 41, Appl
	32	21.4	0.9	60	1	US-08-143-219-18	Sequence 18, Appl
	33	21.2	0.9	54	4	US-08-584-040-8235	Sequence 8235, Ap
	34	21.2	0.9	56	4	US-09-508-542-13	Sequence 13, Appl
c	35	21.2	0.9	58	2	US-08-417-210A-103	Sequence 103, App
	36	21.2	0.9	58	2	US-08-417-210A-108	Sequence 108, App
	37	21	0.9	59	4	US-09-532-656-17	Sequence 17, Appl
	38	21	0.9	59	4	US-09-775-319-4	Sequence 4, Appli
	39	21	0.9	59	4	US-09-532-419A-4	Sequence 4, Appli
	40	21	0.9	60	2	US-08-475-081-2	Sequence 2, Appli
	41	20.6	0.9	47	4	US-09-641-638-1114	Sequence 1114, Ap
	42	20.6	0.9	50	4	US-08-860-038-17	Sequence 17, Appl
	43	20.6	0.9	50	4	US-09-580-923-17	Sequence 17, Appl
	44	20.6	0.9	51	2	US-08-929-501-24	Sequence 24, Appl
	45	20.6	0.9	51	3	US-09-140-177-24	Sequence 24, Appl
	46	20.6	0.9	51	4	US-09-397-979-24	Sequence 24, Appl
	47	20.2	0.9	48	1	US-08-171-389-19	Sequence 19, Appl
	48	20.2	0.9	48	1	US-08-123-936-19	Sequence 19, Appl
	49	20.2	0.9	48	2	US-08-475-228A-19	Sequence 19, Appl
	50	20.2	0.9	48	3	US-08-482-080A-19	Sequence 19, Appl
	51	20.2	0.9	48	4	US-09-354-947-19	Sequence 19, Appl
	52	20.2	0.9	48	5	PCT-US93-12388-19	Sequence 19, Appl
c	53	20.2	0.9	50	1	US-08-171-382-11	Sequence 11, Appl
c	54	20.2	0.9	60	4	US-08-584-760A-33	Sequence 33, Appl
	55	20	0.9	46	1	US-08-222-177A-349	Sequence 349, App
	56	20	0.9	47	4	US-09-338-907-315	Sequence 315, App
	57	20	0.9	47	4	US-09-218-207-315	Sequence 315, App
c	58	20	0.9	49	1	US-08-105-483-102	Sequence 102, App
c	59	20	0.9	49	1	US-08-709-209-102	Sequence 102, App
c	60	20	0.9	49	1	US-08-458-101-102	Sequence 102, App
	61	20	0.9	60	4	US-09-457-959-8	Sequence 8, Appli
	62	19.8	0.8	48	2	US-08-750-128-8	Sequence 8, Appli
	63	19.8	0.8	48	4	US-09-461-697-233	Sequence 233, App
	64	19.8	0.8	49	1	US-08-105-483-103	Sequence 103, App
	65	19.8	0.8	49	1	US-08-709-209-103	Sequence 103, App
	66	19.8	0.8	49	1	US-08-458-101-103	Sequence 103, App
c	67	19.8	0.8	54	4	US-09-423-744A-13	Sequence 13, Appl
c	68	19.8	0.8	58	4	US-09-415-784-11	Sequence 11, Appl
c	69	19.8	0.8	58	4	US-09-415-785A-11	Sequence 11, Appl
c	70	19.8	0.8	58	4	US-08-944-465-11	Sequence 11, Appl
c	71	19.8	0.8	58	4	US-09-415-868-11	Sequence 11, Appl
c	72	19.8	0.8	58	4	US-09-415-900-11	Sequence 11, Appl
c	73	19.6	0.8	48	2	US-08-350-260A-250	Sequence 250, App
	74	19.6	0.8	52	4	US-08-707-860D-3	Sequence 3, Appli
	75	19.6	0.8	57	2	US-08-750-128-11	Sequence 11, Appl
	76	19.6	0.8	60	2	US-08-508-088-1	Sequence 1, Appli
	77	19.6	0.8	60	2	US-09-009-925-1	Sequence 1, Appli
c	78	19.4	0.8	39	1	US-08-068-747-9	Sequence 9, Appli
	79	19.4	0.8	42	3	US-08-938-830-30	Sequence 30, Appl
c	80	19.4	0.8	45	1	US-08-294-770A-6	Sequence 6, Appli
c	81	19.4	0.8	45	2	US-08-448-735C-6	Sequence 6, Appli
	82	19.4	0.8	47	4	US-09-641-638-1045	Sequence 1045, Ap
c	83	19.4	0.8	50	3	US-08-983-607-5	Sequence 5, Appli
c	84	19.4	0.8	50	3	US-08-983-607-42	Sequence 42, Appl
	85	19.4	0.8	55	3	US-09-264-737-4	Sequence 4, Appli
c	86	19.4	0.8	60	1	US-08-592-406-9	Sequence 9, Appli
	87	19.2	0.8	34	1	US-08-741-881-44	Sequence 44, Appl
	88	19.2	0.8	34	1	US-08-739-158-44	Sequence 44, Appl
	89	19.2	0.8	34	2	US-08-739-167-44	Sequence 44, Appl
	90	19.2	0.8	34	3	US-08-404-796-44	Sequence 44, Appl
	91	19.2	0.8	34	3	US-08-931-869-44	Sequence 44, Appl
	92	19.2	0.8	34	4	US-09-350-399-44	Sequence 44, Appl
	93	19.2	0.8	34	4	US-09-236-140A-44	Sequence 44, Appl
	94	19.2	0.8	34	4	US-09-415-784-41	Sequence 41, Appl
	95	19.2	0.8	34	4	US-09-415-785A-41	Sequence 41, Appl
	96	19.2	0.8	34	4	US-08-944-465-41	Sequence 41, Appl
	97	19.2	0.8	34	4	US-09-415-868-41	Sequence 41, Appl
	98	19.2	0.8	34	4	US-09-415-900-41	Sequence 41, Appl
	99	19.2	0.8	40	2	US-09-076-193-1	Sequence 1, Appli
c	100	19.2	0.8	40	4	US-09-605-192-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT 1
US-09-461-697-231
; Sequence 231, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
;  APPLICANT: COGENT NEUROSCIENCE, Inc.
;  APPLICANT: Lo, Donald C.
;  APPLICANT: Barney, Shawn
;  APPLICANT: Thomas, Mary Beth
;  APPLICANT: Portbury, Stuart D.
;  APPLICANT: Puranam, Kasturi
;  APPLICANT: Katz, Lawrence C.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
;  TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
;  TITLE OF INVENTION: CELL DEATH
;  FILE REFERENCE: 10001-005-999
;  CURRENT APPLICATION NUMBER: US/09/461,697
;  CURRENT FILING DATE: 1999-12-14
;  NUMBER OF SEQ ID NOS: 466
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 231
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-461-697-231

  Query Match             1.0%;  Score 23.6;  DB 4;  Length 60;
  Best Local Similarity   64.8%;  Pred. No. 2.1e+03;
  Matches   35;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy   1635 AAAAAGGTGGAGAGGGAGATGTGTGAGCATCCTGCAAGGTGAAACAAGACTCAA 1688
          || ||| || ||  ||| ||| |  || | |  | || | |||| |||| | ||
Db      7 AAGAAGATGAAGGTGGAAATGAGGAAGAAGCTGGAAAAGAGAAAGAAGATTTAA 60


RESULT 2
US-08-198-094-72
; Sequence 72, Application US/08198094
; Patent No. 5741696
;  GENERAL INFORMATION:
;    APPLICANT:  Cochran Ph.D., Mark D
;    TITLE OF INVENTION:  Recombinant Equine Herpesvirus
;    NUMBER OF SEQUENCES:  77
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  John P. White
;      STREET:  30 Rockefeller Plaza
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10112
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/198,094
;      FILING DATE:  February 17, 1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)977-9550
;      TELEFAX:  (212)664-0525
;      TELEX:  422523
;  INFORMATION FOR SEQ ID NO:  72:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  54 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Equine Herpesvirus
US-08-198-094-72

  Query Match             1.0%;  Score 23.4;  DB 1;  Length 54;
  Best Local Similarity   67.3%;  Pred. No. 2.3e+03;
  Matches   33;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    760 CATGATGGAACAGTCACACACACATTTTGTGGAACAATAGAATACATGG 808
          | | ||| | ||  |||| ||    ||  |||  || ||||||||| ||
Db      1 CCTTATGTATCATACACATACGATTTAGGTGACACTATAGAATACACGG 49


RESULT 3
PCT-US95-02087-72
; Sequence 72, Application PC/TUS9502087
;  GENERAL INFORMATION:
;    APPLICANT:  Cochran, Mark D
;    TITLE OF INVENTION:  Recombinant Equine Herpesviruses
;    NUMBER OF SEQUENCES:  77
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  John P. White
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/02087
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/198,094
;      FILING DATE:  February 17, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P
;      REGISTRATION NUMBER:  28,678
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)278-0400
;      TELEFAX:  (212)391-0525
;      TELEX:  422523
;  INFORMATION FOR SEQ ID NO:  72:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  54 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Equine Herpesvirus
PCT-US95-02087-72

  Query Match             1.0%;  Score 23.4;  DB 5;  Length 54;
  Best Local Similarity   67.3%;  Pred. No. 2.3e+03;
  Matches   33;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    760 CATGATGGAACAGTCACACACACATTTTGTGGAACAATAGAATACATGG 808
          | | ||| | ||  |||| ||    ||  |||  || ||||||||| ||
Db      1 CCTTATGTATCATACACATACGATTTAGGTGACACTATAGAATACACGG 49
```

```
RESULT 4
US-08-860-038-15
; Sequence 15, Application US/08860038
; Patent No. 6287762
;  GENERAL INFORMATION:
;    APPLICANT:  CROUZET, Joel
;    APPLICANT:  SCHERMAN, Daniel
;    APPLICANT:  WILS, Pierre
;    TITLE OF INVENTION:  PURIFICATION OF A TRIPLE HELIX FORMATION
;    TITLE OF INVENTION:  WITH AN IMMOBILIZED OLIGONUCLEOTIDE
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rhone-Poulenc Rorer Inc.
;      STREET:  500 Arcola Road, Mailstop 3C43
;      CITY:  Collegeville
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19426
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/860,038
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94/15162
;      FILING DATE:  16-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO FR95/01468
;      FILING DATE:  08-NOV-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Savitzky Esq., Martin F.
;      REGISTRATION NUMBER:  29,699
;      REFERENCE/DOCKET NUMBER:  ST94090-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610) 454-3816
;      TELEFAX:  (610) 454-3808
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  58 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "Oligonucleotide"
US-08-860-038-15

  Query Match             1.0%;  Score 23.4;  DB 4;  Length 58;
  Best Local Similarity   63.2%;  Pred. No. 2.3e+03;
  Matches   36;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy     118 GACCTGGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAG 174
           || | |||  ||   ||||| | ||     ||||| |||  |||||||| ||   ||
Db       1 GATCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 57


RESULT 5
US-09-580-923-15
; Sequence 15, Application US/09580923
; Patent No. 6319672
; GENERAL INFORMATION:
;  APPLICANT:  Crouzet, Joel
;  APPLICANT:  Scherman, Daniel
;  APPLICANT:  Wils, Pierre
;  APPLICANT:  Cameron, Beatrice
;  APPLICANT:  Blanche, Francis
;  TITLE OF INVENTION:  PURIFICATION OF A TRIPLE HELIX FORMATION WITH AN
;  TITLE OF INVENTION:  IMMOBILIZED OLIGONUCLEOTIDE
;  FILE REFERENCE:  03804.0138-01
;  CURRENT APPLICATION NUMBER: US/09/580,923
;  CURRENT FILING DATE:  2000-05-26
;  PRIOR APPLICATION NUMBER:  08/860,038
;  PRIOR FILING DATE:  1997-06-09
;  PRIOR APPLICATION NUMBER:  PCT/FR95/01468
;  PRIOR FILING DATE:  1995-11-08
;  NUMBER OF SEQ ID NOS:  36
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 58
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: oligonucleotide
US-09-580-923-15

  Query Match             1.0%;  Score 23.4;  DB 4;  Length 58;
  Best Local Similarity   63.2%;  Pred. No. 2.3e+03;
  Matches   36;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy     118 GACCTGGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAG 174
           || | |||  ||   ||||| | ||     ||||| |||  |||||||| ||   ||
Db       1 GATCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 57


RESULT 6
US-08-823-516-162
; Sequence 162, Application US/08823516
; Patent No. 5994069
;  GENERAL INFORMATION:
;    APPLICANT:  Hall, Jeff G.
;    APPLICANT:  Lyamichev, Victor I.
;    APPLICANT:  Mast, Andrea L.
;    APPLICANT:  Brow, Mary Ann D.
;    TITLE OF INVENTION:  Detection Of Nucleic Acids By Multiple
;    TITLE OF INVENTION:  Sequential Invasive Cleavages
;    NUMBER OF SEQUENCES:  163
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Medlen & Carroll, LLP
;      STREET:  220 Montgomery Street, Suite 2200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States Of America
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/823,516
;      FILING DATE:  24-MAR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US97/01072
;      FILING DATE:  21-JAN-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/759,038
;      FILING DATE:  02-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/758,314
;      FILING DATE:  02-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/756,386
;      FILING DATE:  29-NOV-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/682,853
;      FILING DATE:  12-JUL-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/599,491
;      FILING DATE:  24-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
```

```
;      NAME:  Ingolia, Diane E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  FORS-02736
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  162:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  54 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "DNA"
US-08-823-516-162

  Query Match             1.0%;  Score 22.6;  DB 2;  Length 54;
  Best Local Similarity   75.7%;  Pred. No. 3.7e+03;
  Matches   28;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy     128 AGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGA 164
           ||  |||||| | |||  || |||| |||| ||||||
Db       4 AGGAAGAGGAGGAGGGTGCTCAGGAGGAGCGGGAGGA 40


RESULT 7
US-08-171-389-63
; Sequence 63, Application US/08171389
; Patent No. 5578444
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/171,389
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175/G19P3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  63:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human sodium/potassium ATPase alpha
;      INDIVIDUAL ISOLATE:  3 subunit (ATP1 A3)
US-08-171-389-63

  Query Match             1.0%;  Score 22.4;  DB 1;  Length 44;
  Best Local Similarity   72.5%;  Pred. No. 3.8e+03;
  Matches   29;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       9 CTGCGGCGGGTCCGGGCCCATGAGGCGACGAAGGAGGCGG 48
           || | ||||   |||||  |||||| | || ||| |||||
Db       4 CTCCCGCGGACGCGGGCATATGAGGAGGCGGAGGCGGCGG 43


RESULT 8
US-08-123-936-63
; Sequence 63, Application US/08123936
; Patent No. 5726014
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    TITLE OF INVENTION:  Screening Assay for the Detection of
;    TITLE OF INVENTION:  DNA-Binding Molecules
;    NUMBER OF SEQUENCES:  640
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/123,936
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0075.32/G19P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  63:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human sodium/potassium ATPase alpha
```

```
;      INDIVIDUAL ISOLATE:  3 subunit (ATP1 A3)
US-08-123-936-63

  Query Match             1.0%;  Score 22.4;  DB 1;  Length 44;
  Best Local Similarity   72.5%;  Pred. No. 3.8e+03;
  Matches   29;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       9 CTGCGGCGGGTCCGGGCCCATGAGGCGACGAAGGAGGCGG 48
           || | ||||   |||||  |||||| | || ||| |||||
Db       4 CTCCCGCGGACGCGGGCATATGAGGAGGCGGAGGCGGCGG 43


RESULT 9
US-08-475-228A-63
; Sequence 63, Application US/08475228A
; Patent No. 5869241
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,228A
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stratford, Carol A.
;      REGISTRATION NUMBER:  34,444
;      REFERENCE/DOCKET NUMBER:  4600-0175.21/G19P3D2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  63:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human sodium/potassium ATPase alpha
;      INDIVIDUAL ISOLATE:  3 subunit (ATP1 A3)
US-08-475-228A-63

  Query Match             1.0%;  Score 22.4;  DB 2;  Length 44;
  Best Local Similarity   72.5%;  Pred. No. 3.8e+03;
  Matches   29;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       9 CTGCGGCGGGTCCGGGCCCATGAGGCGACGAAGGAGGCGG 48
           || | ||||   |||||  |||||| | || ||| |||||
Db       4 CTCCCGCGGACGCGGGCATATGAGGAGGCGGAGGCGGCGG 43


RESULT 10
US-08-482-080A-63
; Sequence 63, Application US/08482080A
; Patent No. 6010849
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,080A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/171,389
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brady, John F.
;      REGISTRATION NUMBER:  39,118
;      REFERENCE/DOCKET NUMBER:  4600-0175.20/G19P3D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 324-0880
;      TELEFAX:  (650) 324-0960
;  INFORMATION FOR SEQ ID NO:  63:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human sodium/potassium ATPase alpha
;      INDIVIDUAL ISOLATE:  3 subunit (ATP1 A3)
US-08-482-080A-63

  Query Match             1.0%;  Score 22.4;  DB 3;  Length 44;
  Best Local Similarity   72.5%;  Pred. No. 3.8e+03;
```

```
  Matches   29;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       9 CTGCGGCGGGTCCGGGCCCATGAGGCGACGAAGGAGGCGG 48
           || | ||||   |||||  |||||| | || ||| |||||
Db       4 CTCCCGCGGACGCGGGCATATGAGGAGGCGGAGGCGGCGG 43


RESULT 11
US-09-354-947-63
; Sequence 63, Application US/09354947
; Patent No. 6384208
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/354,947
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/482,080
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/171,389
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brady, John F.
;      REGISTRATION NUMBER:  39,118
;      REFERENCE/DOCKET NUMBER:  4600-0175.20/G19P3D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 324-0880
;      TELEFAX:  (650) 324-0960
;  INFORMATION FOR SEQ ID NO:  63:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human sodium/potassium ATPase alpha
;      INDIVIDUAL ISOLATE:  3 subunit (ATP1 A3)
US-09-354-947-63

  Query Match              1.0%;  Score 22.4;  DB 4;  Length 44;
  Best Local Similarity   72.5%;  Pred. No. 3.8e+03;
  Matches   29;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       9 CTGCGGCGGGTCCGGGCCCATGAGGCGACGAAGGAGGCGG 48
           || | ||||   |||||  |||||| | || ||| |||||
Db       4 CTCCCGCGGACGCGGGCATATGAGGAGGCGGAGGCGGCGG 43


RESULT 12
PCT-US93-12388-63
; Sequence 63, Application PC/TUS9312388
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12388
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175.41/G19PCT2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  63:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human sodium/potassium ATPase alpha
;      INDIVIDUAL ISOLATE:  3 subunit (ATP1 A3)
PCT-US93-12388-63

  Query Match              1.0%;  Score 22.4;  DB 5;  Length 44;
  Best Local Similarity   72.5%;  Pred. No. 3.8e+03;
  Matches   29;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       9 CTGCGGCGGGTCCGGGCCCATGAGGCGACGAAGGAGGCGG 48
           || | ||||   |||||  |||||| | || ||| |||||
Db       4 CTCCCGCGGACGCGGGCATATGAGGAGGCGGAGGCGGCGG 43


RESULT 13
US-08-171-389-212
; Sequence 212, Application US/08171389
; Patent No. 5578444
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
```

```
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/171,389
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175/G19P3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  212:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human interleukin 4 gene
US-08-171-389-212

  Query Match             0.9%;  Score 22.2;  DB 1;  Length 45;
  Best Local Similarity   77.1%;  Pred. No. 4.4e+03;
  Matches   27;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    933 TGACAAAATCCTCAAATGTAAACTCAATTTGCCTC 967
          | || |||   || ||||||||||||| ||| ||||
Db      9 TAACGAAAATTTCCAATGTAAACTCATTTTCCCTC 43


RESULT 14
US-08-123-936-212
; Sequence 212, Application US/08123936
; Patent No. 5726014
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    TITLE OF INVENTION:  Screening Assay for the Detection of
;    TITLE OF INVENTION:  DNA-Binding Molecules
;    NUMBER OF SEQUENCES:  640
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/123,936
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0075.32/G19P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  212:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human interleukin 4 gene
US-08-123-936-212

  Query Match             0.9%;  Score 22.2;  DB 1;  Length 45;
  Best Local Similarity   77.1%;  Pred. No. 4.4e+03;
  Matches   27;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    933 TGACAAAATCCTCAAATGTAAACTCAATTTGCCTC 967
          | || |||   || ||||||||||||| ||| ||||
Db      9 TAACGAAAATTTCCAATGTAAACTCATTTTCCCTC 43


RESULT 15
US-08-475-228A-212
; Sequence 212, Application US/08475228A
; Patent No. 5869241
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/475,228A
;       FILING DATE:  06-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/123,936
;       FILING DATE:  17-SEP-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/996,783
;       FILING DATE:  23-DEC-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/723,618
;       FILING DATE:  27-JUN-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/081,070
;       FILING DATE:  22-JUN-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Stratford, Carol A.
;       REGISTRATION NUMBER:  34,444
;       REFERENCE/DOCKET NUMBER:  4600-0175.21/G19P3D2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 324-0880
;       TELEFAX:  (415) 324-0960
;   INFORMATION FOR SEQ ID NO:  212:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  45 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ORIGINAL SOURCE:
;       INDIVIDUAL ISOLATE:  Human interleukin 4 gene
US-08-475-228A-212

  Query Match             0.9%;  Score 22.2;  DB 2;  Length 45;
  Best Local Similarity   77.1%;  Pred. No. 4.4e+03;
  Matches   27;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     933 TGACAAAATCCTCAAATGTAAACTCAATTTGCCTC 967
           | || |||   || ||||||||||||| ||| ||||
Db       9 TAACGAAAATTTCCAATGTAAACTCATTTTCCCTC 43


RESULT 16
US-08-482-080A-212
; Sequence 212, Application US/08482080A
; Patent No. 6010849
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,080A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/171,389
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brady, John F.
;      REGISTRATION NUMBER:  39,118
;      REFERENCE/DOCKET NUMBER:  4600-0175.20/G19P3D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 324-0880
;      TELEFAX:  (650) 324-0960
;  INFORMATION FOR SEQ ID NO:  212:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human interleukin 4 gene
US-08-482-080A-212

  Query Match             0.9%;  Score 22.2;  DB 3;  Length 45;
  Best Local Similarity   77.1%;  Pred. No. 4.4e+03;
  Matches   27;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     933 TGACAAAATCCTCAAATGTAAACTCAATTTGCCTC 967
           | || |||   || ||||||||||||| ||| ||||
Db       9 TAACGAAAATTTCCAATGTAAACTCATTTTCCCTC 43


RESULT 17
US-09-354-947-212
; Sequence 212, Application US/09354947
; Patent No. 6384208
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/354,947
;      FILING DATE:
```

```
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/482,080
;       FILING DATE:  07-JUN-1995
;       APPLICATION NUMBER:  US 08/171,389
;       FILING DATE:  20-DEC-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/123,936
;       FILING DATE:  17-SEP-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/996,783
;       FILING DATE:  23-DEC-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/723,618
;       FILING DATE:  27-JUN-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/081,070
;       FILING DATE:  22-JUN-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Brady, John F.
;       REGISTRATION NUMBER:  39,118
;       REFERENCE/DOCKET NUMBER:  4600-0175.20/G19P3D1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (650) 324-0880
;       TELEFAX:  (650) 324-0960
;   INFORMATION FOR SEQ ID NO:  212:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  45 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ORIGINAL SOURCE:
;       INDIVIDUAL ISOLATE:  Human interleukin 4 gene
US-09-354-947-212

  Query Match             0.9%;  Score 22.2;  DB 4;  Length 45;
  Best Local Similarity   77.1%;  Pred. No. 4.4e+03;
  Matches   27;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     933 TGACAAAATCCTCAAATGTAAACTCAATTTGCCTC 967
           | || |||   || ||||||||||| ||| ||||
Db       9 TAACGAAAATTTCCAATGTAAACTCATTTTCCCTC 43


RESULT 18
PCT-US93-12388-212
; Sequence 212, Application PC/TUS9312388
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12388
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175.41/G19PCT2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  212:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human interleukin 4 gene
PCT-US93-12388-212

  Query Match             0.9%;  Score 22.2;  DB 5;  Length 45;
  Best Local Similarity   77.1%;  Pred. No. 4.4e+03;
  Matches   27;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     933 TGACAAAATCCTCAAATGTAAACTCAATTTGCCTC 967
           | || |||   || ||||||||||| ||| ||||
Db       9 TAACGAAAATTTCCAATGTAAACTCATTTTCCCTC 43


RESULT 19
US-09-641-638-880/c
; Sequence 880, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Cohen, Annick
;  TITLE OF INVENTION:  BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
;  TITLE OF INVENTION:  GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
;  FILE REFERENCE:  GENSET.051CP1
;  CURRENT APPLICATION NUMBER: US/09/641,638
;  CURRENT FILING DATE:  2000-08-16
;  PRIOR APPLICATION NUMBER:  US 09/502,330
;  PRIOR FILING DATE:  2000-02-11
;  PRIOR APPLICATION NUMBER:  US 60/133,200
;  PRIOR FILING DATE:  1999-05-07
;  PRIOR APPLICATION NUMBER:  US 09/275,267
;  PRIOR FILING DATE:  1999-03-23
;  PRIOR APPLICATION NUMBER:  US 60/119,917
;  PRIOR FILING DATE:  1999-02-12
;  NUMBER OF SEQ ID NOS:  1304
;  SOFTWARE:  Patent.pm
; SEQ ID NO 880
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 10-88-81 : polymorphic base C or T
US-09-641-638-880

  Query Match             0.9%;  Score 22.2;  DB 4;  Length 47;
  Best Local Similarity   66.7%;  Pred. No. 4.5e+03;
  Matches   30;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    2107 ATCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGAT 2151
           || |||  || || | || ||| :|||| ||  || | |||| ||
Db      47 ATGTTTGGTTGATTTCAACATAGRTTTTCCAGTTATACTTTTAAT 3
```

```
RESULT 20
US-08-975-703-31
; Sequence 31, Application US/08975703
; Patent No. 6030832
;  GENERAL INFORMATION:
;    APPLICANT:  Wong, Alexander K.C.
;    APPLICANT:  Bartel, Paul L.
;    APPLICANT:  Teng, David H.-F.
;    APPLICANT:  Tavtigian, Sean V.
;    TITLE OF INVENTION:  A Carboxy-Terminal BRCA1 Interacting
;    TITLE OF INVENTION:  Protein
;    NUMBER OF SEQUENCES:  41
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rothwell, Figg, Ernst & Kurz, P.C.
;      STREET:  555 Thirteenth Street, N.W., Suite 701 East
;      STREET:  Tower
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.A.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/975,703
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Saxe, Stephen A.
;      REGISTRATION NUMBER:  38,609
;      REFERENCE/DOCKET NUMBER:  2318-0174
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-624-1589
;      TELEFAX:  202-783-6031
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  42 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "Primer"
US-08-975-703-31

  Query Match             0.9%;  Score 22;  DB 3;  Length 42;
  Best Local Similarity   73.7%;  Pred. No. 4.8e+03;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   2034 TTTTCCCAGGCAAAATGGGAAGGCAAAGACAAAAGAAA 2071
          ||||||||| ||  | ||||||| ||  |||| | | |
Db      2 TTTTCCCAGTCACGACGGGAAGGAAATCACAAGAAACA 39


RESULT 21
US-09-515-884-31
; Sequence 31, Application US/09515884
; Patent No. 6235263
;   GENERAL INFORMATION:
;        APPLICANT: Wong, Alexander K.C.
;                   Bartel, Paul L.
;                   Teng, David H.-F.
;                   Tavtigian, Sean V.
;        TITLE OF INVENTION: A Carboxy-Terminal BRCA1 Interacting
;                            Protein
;        NUMBER OF SEQUENCES: 41
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
;             STREET: 555 Thirteenth Street, N.W., Suite 701 East
;                     Tower
;             CITY: Washington
;             STATE: DC
;             COUNTRY: U.S.A.
;             ZIP: 20004
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/515,884
;             FILING DATE: 29-Feb-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/975,703
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Saxe, Stephen A.
;             REGISTRATION NUMBER: 38,609
;             REFERENCE/DOCKET NUMBER: 2318-0174
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 202-624-1589
;             TELEFAX: 202-783-6031
;   INFORMATION FOR SEQ ID NO: 31:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 42 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: other nucleic acid
;             DESCRIPTION: /desc = "Primer"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-515-884-31

  Query Match             0.9%;  Score 22;  DB 4;  Length 42;
  Best Local Similarity   73.7%;  Pred. No. 4.8e+03;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   2034 TTTTCCCAGGCAAAATGGGAAGGCAAAGACAAAAGAAA 2071
          ||||||||| ||  | ||||||| ||  |||| | | |
Db      2 TTTTCCCAGTCACGACGGGAAGGAAATCACAAGAAACA 39


RESULT 22
US-08-860-038-16/c
; Sequence 16, Application US/08860038
; Patent No. 6287762
;  GENERAL INFORMATION:
;    APPLICANT:  CROUZET, Joel
;    APPLICANT:  SCHERMAN, Daniel
;    APPLICANT:  WILS, Pierre
;    TITLE OF INVENTION:  PURIFICATION OF A TRIPLE HELIX FORMATION
;    TITLE OF INVENTION:  WITH AN IMMOBILIZED OLIGONUCLEOTIDE
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rhone-Poulenc Rorer Inc.
;      STREET:  500 Arcola Road, Mailstop 3C43
;      CITY:  Collegeville
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19426
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/860,038
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94/15162
;      FILING DATE:  16-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO FR95/01468
```

```
;       FILING DATE:  08-NOV-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Savitzky Esq., Martin F.
;       REGISTRATION NUMBER:  29,699
;       REFERENCE/DOCKET NUMBER:  ST94090-US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (610) 454-3816
;       TELEFAX:  (610) 454-3808
;   INFORMATION FOR SEQ ID NO:  16:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  58 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  other nucleic acid
;       DESCRIPTION:  /desc = "Oligonucleotide"
US-08-860-038-16

  Query Match             0.9%;  Score 22;  DB 4;  Length 58;
  Best Local Similarity   63.0%;  Pred. No. 5.7e+03;
  Matches   34;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy     123 GGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAGTT 176
           |||  ||   ||||| | ||     ||||| |||  |||||||| ||   | ||
Db      54 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAATT 1


RESULT 23
US-09-580-923-16/c
; Sequence 16, Application US/09580923
; Patent No. 6319672
; GENERAL INFORMATION:
;  APPLICANT:  Crouzet, Joel
;  APPLICANT:  Scherman, Daniel
;  APPLICANT:  Wils, Pierre
;  APPLICANT:  Cameron, Beatrice
;  APPLICANT:  Blanche, Francis
;  TITLE OF INVENTION:  PURIFICATION OF A TRIPLE HELIX FORMATION WITH AN
;  TITLE OF INVENTION:  IMMOBILIZED OLIGONUCLEOTIDE
;  FILE REFERENCE:  03804.0138-01
;  CURRENT APPLICATION NUMBER: US/09/580,923
;  CURRENT FILING DATE:  2000-05-26
;  PRIOR APPLICATION NUMBER:  08/860,038
;  PRIOR FILING DATE:  1997-06-09
;  PRIOR APPLICATION NUMBER:  PCT/FR95/01468
;  PRIOR FILING DATE:  1995-11-08
;  NUMBER OF SEQ ID NOS:  36
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 16
;   LENGTH: 58
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: oligonucleotide
US-09-580-923-16

  Query Match             0.9%;  Score 22;  DB 4;  Length 58;
  Best Local Similarity   63.0%;  Pred. No. 5.7e+03;
  Matches   34;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy     123 GGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAGTT 176
           |||  ||   ||||| | ||     ||||| |||  |||||||| ||   | ||
Db      54 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAATT 1


RESULT 24
US-09-021-560-5
; Sequence 5, Application US/09021560
; Patent No. 6410719
;  GENERAL INFORMATION:
;    APPLICANT:  BOREN, THOMAS
;    APPLICANT:  NORMARK, STAFFAN
;    APPLICANT:  ARNQVIST, ANNA
;    APPLICANT:  ILVER, DAG
;    TITLE OF INVENTION:  BLOOD GROUP ANTIGEN BINDING PROTEIN AND
;    TITLE OF INVENTION:  CORRESPONDING GENES
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH AND BIRCH
;      STREET:  PO BOX 747
;      CITY:  FALLS CHURCH
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22040-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/021,560
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURPHY JR, GERALD M
;      REGISTRATION NUMBER:  28,977
;      REFERENCE/DOCKET NUMBER:  825-144P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  60 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
US-09-021-560-5

  Query Match             0.9%;  Score 22;  DB 4;  Length 60;
  Best Local Similarity   63.0%;  Pred. No. 5.8e+03;
  Matches   34;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy      49 GACGGCTTTTACCCAGCCCCGGACTTCCGAGACAGGGAAGCTGAGGACATGGCA 102
           ||||||||||||  ||  |   | ||  | |  | | ||||| |   | |||| |
Db       7 GACGGCTTTTACACAAGCGTAGGCTATCAAATCGGTGAAGCCGCTCAAATGGTA 60


RESULT 25
US-08-983-607-7/c
; Sequence 7, Application US/08983607
; Patent No. 6140470
;  GENERAL INFORMATION:
;    APPLICANT:  Alan Garen
;    APPLICANT:  Xiaohong Cai
;    TITLE OF INVENTION:  Human Anti-Tumor Monoclonal Anti-
;    TITLE OF INVENTION:  bodies
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Department of Molecular Biophysics
;      ADDRESSEE:  and Biochemistry, Yale University
;      STREET:  266 Whitney Avenue
;      CITY:  New Haven
;      STATE:  Connecticut
;      COUNTRY:  United States of America
;      ZIP:  06520-8114
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" 1.44 Mb diskette
;      COMPUTER:  IBM PC
;      OPERATING SYSTEM:  MS DOS
;      SOFTWARE:  Word Processing
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/983,607
;      FILING DATE:  April 27, 1998
;      CLASSIFICATION:  435
```

```
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/IB96/01032
;       FILING DATE:  June 28, 1996
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Mary M. Krinsky
;       REGISTRATION NUMBER:  32423
;       REFERENCE/DOCKET NUMBER:  OCR-679
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  203-773-9544
;       TELEFAX:  203-773-1183
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  50 residues
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
;       DESCRIPTION:  primer used in constructs
US-08-983-607-7

  Query Match             0.9%;  Score 21.8;  DB 3;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 5.9e+03;
  Matches   29;  Conservative    1;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    1501 GCATCGGCACCACTTCCAATACGACAGCCGAACTCTGGGCCAT 1543
           :|| | ||  ||| ||    ||| | ||||||| ||| |||||
Db      43 SCAMCAGCTGCACCTCGGCCACGTCGGCCGAACCCTGAGCCAT 1


RESULT 26
US-08-983-607-44/c
; Sequence 44, Application US/08983607
; Patent No. 6140470
;  GENERAL INFORMATION:
;    APPLICANT:  Alan Garen
;    APPLICANT:  Xiaohong Cai
;    TITLE OF INVENTION:  Human Anti-Tumor Monoclonal Anti-
;    TITLE OF INVENTION:  bodies
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Department of Molecular Biophysics
;      ADDRESSEE:  and Biochemistry, Yale University
;      STREET:  266 Whitney Avenue
;      CITY:  New Haven
;      STATE:  Connecticut
;      COUNTRY:  United States of America
;      ZIP:  06520-8114
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" 1.44 Mb diskette
;      COMPUTER:  IBM PC
;      OPERATING SYSTEM:  MS DOS
;      SOFTWARE:  Word Processing
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/983,607
;      FILING DATE:  April 27, 1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/IB96/01032
;      FILING DATE:  June 28, 1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary M. Krinsky
;      REGISTRATION NUMBER:  32423
;      REFERENCE/DOCKET NUMBER:  OCR-679
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  203-773-9544
;      TELEFAX:  203-773-1183
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 residues
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;      DESCRIPTION:  primer used in constructs
US-08-983-607-44

  Query Match             0.9%;  Score 21.8;  DB 3;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 5.9e+03;
  Matches   29;  Conservative    1;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    1501 GCATCGGCACCACTTCCAATACGACAGCCGAACTCTGGGCCAT 1543
           :|| | ||  ||| ||    ||| | ||||||| ||| |||||
Db      43 SCAMCAGCTGCACCTCGGCCACGTCGGCCGAACCCTGAGCCAT 1


RESULT 27
US-08-821-559A-12
; Sequence 12, Application US/08821559A
; Patent No. 5846774
;  GENERAL INFORMATION:
;    APPLICANT:  XIA, YUANNAN
;    TITLE OF INVENTION:  CHLORELLA VIRUS PROMOTERS
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merchant, Gould, Smith, Edell, Welter & Schmidt
;      STREET:  3100 No. 5846774west Center, 90 South Seventh St
;      CITY:  Minneapolis
;      STATE:  MN
;      COUNTRY:  USA
;      ZIP:  55402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/821,559A
;      FILING DATE:  21-MAR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kettelberger, Denise M
;      REGISTRATION NUMBER:  33,924
;      REFERENCE/DOCKET NUMBER:  8648.63-US-01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  612/371-5268
;      TELEFAX:  612/332-9081
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-821-559A-12

  Query Match             0.9%;  Score 21.4;  DB 2;  Length 48;
  Best Local Similarity   66.0%;  Pred. No. 7.5e+03;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    2233 CTTCTCTTGGTGAAATAATAAAATGCAAATGAATCATTGTTAACACA 2279
           ||| || || ||| || |||||   |||||||  |  |||    |||
Db       2 CTTATCGTGTTGATATGATAAATGACAAATGATACGCTGTATCAACA 48


RESULT 28
US-09-400-541-12
; Sequence 12, Application US/09400541
; Patent No. 6316224
;   GENERAL INFORMATION:
;        APPLICANT: BIONEBRASKA, INC.
```

```
;       TITLE OF INVENTION: CHLORELLA VIRUS PROMOTERS
;       NUMBER OF SEQUENCES: 25
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
;            STREET: 3100 No. 6316224west Center, 90 South 7th Street
;            CITY: Minneapolis
;            STATE: MN
;            COUNTRY: USA
;            ZIP: 55402
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/400,541
;            FILING DATE: 21-Sep-1999
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/821,559
;            FILING DATE: 21-MAR-1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Kettelberger, Denise M
;            REGISTRATION NUMBER: 33,924
;            REFERENCE/DOCKET NUMBER: 8648.63WOI1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 612-332-5300
;            TELEFAX: 612-332-9081
;            TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 12:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 48 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-400-541-12

  Query Match             0.9%;  Score 21.4;  DB 4;  Length 48;
  Best Local Similarity   66.0%;  Pred. No. 7.5e+03;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   2233 CTTCTCTTGGTGAAATAATAAAATGCAAATGAATCATTGTTAACACA 2279
          ||| || || ||| || |||||   |||||||  |  |||    |||
Db      2 CTTATCGTGTTGATATGATAAATGACAAATGATACGCTGTATCAACA 48


RESULT 29
US-09-580-923-34
; Sequence 34, Application US/09580923
; Patent No. 6319672
; GENERAL INFORMATION:
;  APPLICANT:  Crouzet, Joel
;  APPLICANT:  Scherman, Daniel
;  APPLICANT:  Wils, Pierre
;  APPLICANT:  Cameron, Beatrice
;  APPLICANT:  Blanche, Francis
;  TITLE OF INVENTION:  PURIFICATION OF A TRIPLE HELIX FORMATION WITH AN
;  TITLE OF INVENTION:  IMMOBILIZED OLIGONUCLEOTIDE
;  FILE REFERENCE:  03804.0138-01
;  CURRENT APPLICATION NUMBER: US/09/580,923
;  CURRENT FILING DATE:  2000-05-26
;  PRIOR APPLICATION NUMBER:  08/860,038
;  PRIOR FILING DATE:  1997-06-09
;  PRIOR APPLICATION NUMBER:  PCT/FR95/01468
;  PRIOR FILING DATE:  1995-11-08
;  NUMBER OF SEQ ID NOS:  36
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 34
;   LENGTH: 48
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: oligonucleotide
US-09-580-923-34

  Query Match             0.9%;  Score 21.4;  DB 4;  Length 48;
  Best Local Similarity   66.0%;  Pred. No. 7.5e+03;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    123 GGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGG 169
          |||  ||   ||||| | ||     ||||| |||  |||||||| ||
Db      1 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 47


RESULT 30
US-09-899-999-12
; Sequence 12, Application US/09899999
; Patent No. 6395965
;   GENERAL INFORMATION:
;       APPLICANT: BIONEBRASKA, INC.
;       TITLE OF INVENTION: CHLORELLA VIRUS PROMOTERS
;       NUMBER OF SEQUENCES: 25
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
;            STREET: 3100 No. 6395965west Center, 90 South 7th Street
;            CITY: Minneapolis
;            STATE: MN
;            COUNTRY: USA
;            ZIP: 55402
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/899,999
;            FILING DATE: 09-Jul-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/400,541
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Kettelberger, Denise M
;            REGISTRATION NUMBER: 33,924
;            REFERENCE/DOCKET NUMBER: 8648.63WOI1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 612-332-5300
;            TELEFAX: 612-332-9081
;            TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 12:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 48 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-899-999-12

  Query Match             0.9%;  Score 21.4;  DB 4;  Length 48;
  Best Local Similarity   66.0%;  Pred. No. 7.5e+03;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   2233 CTTCTCTTGGTGAAATAATAAAATGCAAATGAATCATTGTTAACACA 2279
          ||| || || ||| || |||||   |||||||  |  |||    |||
Db      2 CTTATCGTGTTGATATGATAAATGACAAATGATACGCTGTATCAACA 48


RESULT 31
US-08-113-646A-41
; Sequence 41, Application US/08113646A
; Patent No. 5578468
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  PICKUP, David J.
;    APPLICANT:  PATEL, Dhavalkumar
;    APPLICANT:  ANTCZAK, James B.
;    TITLE OF INVENTION:  SITE-SPECIFIC RNA CLEAVAGE
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 NORTH GLEBE ROAD
;      CITY:  ARLINGTON
;      STATE:  VIRGINIA
;      COUNTRY:  U.S.A.
;      ZIP:  22201-4714
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/113,646A
;      FILING DATE:  31-AUG-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/084,406
;      FILING DATE:  10-AUG-1987
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WILSON, MARY J.
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  1579-20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 816-4000
;      TELEFAX:  (703) 816-4100
;      TELEX:  200797 NIXN UR
;  INFORMATION FOR SEQ ID NO:  41:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  55 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA (genomic)
US-08-113-646A-41

  Query Match             0.9%;  Score 21.4;  DB 1;  Length 55;
  Best Local Similarity   57.4%;  Pred. No. 8.1e+03;
  Matches   27;  Conservative    4;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   1754 GATTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAA 1800
          ||:::| ||||  ||  ||      |||||||||   |||  ||||
Db      8 GAUUUUAUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 54


RESULT 32
US-08-143-219-18
; Sequence 18, Application US/08143219
; Patent No. 5670330
;  GENERAL INFORMATION:
;    APPLICANT:  Sonenberg, Nahum
;    APPLICANT:  Katze, Michael G.
;    APPLICANT:  Roy, Sophie
;    APPLICANT:  Koromilas, Antonis E.
;    APPLICANT:  Barber, Glen N.
;    TITLE OF INVENTION:  TUMOR-CELL ASSAY METHOD AND KIT
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  611 West Sixth Street
;      CITY:  Los Angeles
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  90017
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  PC-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/143,219
;      FILING DATE:  October 25, 1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:                    two
;      APPLICATION NUMBER:  08/141,244
;      FILING DATE:  October 22, 1993
;      APPLICATION NUMBER:  07/953,681
;      FILING DATE:  September 29, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Douglas E. Olson
;      REGISTRATION NUMBER:  22,798
;      REFERENCE/DOCKET NUMBER:  204/139
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  60 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  COMPLEMENTARY TO THE RNA PROBE FOR
;      INDIVIDUAL ISOLATE:  PR-VI, FIGURE 5
US-08-143-219-18

  Query Match             0.9%;  Score 21.4;  DB 1;  Length 60;
  Best Local Similarity   61.8%;  Pred. No. 8.4e+03;
  Matches   34;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy    644 TGGGGCATTTACATCAAAAGGGGATCATCTACAGAGACCTGAAGCCGGAGAATAT 698
          |||   || |||||  |||     | ||  | ||||| || |||||    |||||
Db      5 TGGATTATATACATTCAAAAAAATTAATTCATAGAGATCTTAAGCCAAGTAATAT 59


RESULT 33
US-08-584-040-8235
; Sequence 8235, Application US/08584040
; Patent No. 6346398
;  GENERAL INFORMATION:
;    APPLICANT:  Pavco, Pamela
;    APPLICANT:  McSwiggen, James
;    APPLICANT:  Stinchcomb, Dan T.
;    APPLICANT:  Escobedo, Jaime
;    TITLE OF INVENTION:  METHOD AND REAGENT FOR THE
;    TITLE OF INVENTION:  TREATMENT OF DISEASES OR
;    TITLE OF INVENTION:  CONDITIONS RELATED TO LEVELS
;    TITLE OF INVENTION:  OF VASCULAR ENDOTHELIAL
;    TITLE OF INVENTION:  GROWTH FACTOR
;    NUMBER OF SEQUENCES:  8502
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  Word Perfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/584,040
```

```
;       FILING DATE:  January 11, 1996
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  60/005,974
;       FILING DATE:  October 26, 1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Warburg, Richard J.
;       REGISTRATION NUMBER:  32,327
;       REFERENCE/DOCKET NUMBER:  218/064
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (213) 489-1600
;       TELEFAX:  (213) 955-0440
;       TELEX:  67-3510
;   INFORMATION FOR SEQ ID NO:  8235:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  54 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-584-040-8235

  Query Match             0.9%;  Score 21.2;  DB 4;  Length 54;
  Best Local Similarity   52.0%;  Pred. No. 9e+03;
  Matches   26;  Conservative    6;  Mismatches   18;  Indels    0;  Gaps    0;

Qy   2045 AAAATGGGAAGGCAAAGACAAAAGAAACTTACCAATTGATGTTTTACGTG 2094
          ||  :||  | |||||||| | ||||||  ||    :| :   ::|| :|
Db      2 AAUGUGGAGAAGCAAAGACCAGAGAAACACACGUUGUGGUACAUUACCUG 51


RESULT 34
US-09-508-542-13
; Sequence 13, Application US/09508542
; Patent No. 6339174
; GENERAL INFORMATION:
;  APPLICANT: STRAUSS, ANDREAS
;  APPLICANT:  THUMM, GUNTHER
;  APPLICANT:  POHLNER, JOHANNES
;  APPLICANT:  GOTZ, FRIEDRICH
;  TITLE OF INVENTION: METHOD FOR IDENTIFYING A NUCLEIC ACID
;  FILE REFERENCE: 10496/P65266US0
;  CURRENT APPLICATION NUMBER: US/09/508,542
;  CURRENT FILING DATE:  2000-05-16
;  PRIOR APPLICATION NUMBER: PCT/EP98/06136
;  PRIOR FILING DATE: 1998-09-26
;  PRIOR APPLICATION NUMBER: 97 116 841.4
;  PRIOR FILING DATE: 1997-09-27
;  PRIOR APPLICATION NUMBER: 97 118 755.4
;  PRIOR FILING DATE: 1997-10-29
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
;   LENGTH: 56
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: oligonucleotide
US-09-508-542-13

  Query Match             0.9%;  Score 21.2;  DB 4;  Length 56;
  Best Local Similarity   69.0%;  Pred. No. 9.2e+03;
  Matches   29;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    221 GCATGGAACATTGTGAGAAATTTGAAATCTCAGAAACTAGTG 262
          || |   ||| | | |||||| |||||| ||  || | ||||
Db      9 GCTTACCACAATCTAAGAAATCTGAAATATCTCAAGCAAGTG 50


RESULT 35
US-08-417-210A-103/c
; Sequence 103, Application US/08417210A
; Patent No. 5863542
;  GENERAL INFORMATION:
;    APPLICANT:  PAOLETTI, ENZO
;    APPLICANT:  TARTAGLIA, JAMES
;    APPLICANT:  COX, WILLIAM I.
;    TITLE OF INVENTION:  IMMUNODEFICIENCY RECOMBINANT POXVIRUS
;    NUMBER OF SEQUENCES:  148
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CURTIS, MORRIS & SAFFORD, P.C.
;      STREET:  530 FIFTH AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/417,210A
;      FILING DATE:  05-APR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KOWALSKI, THOMAS J.
;      REGISTRATION NUMBER:  32,147
;      REFERENCE/DOCKET NUMBER:  454310-2690
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-840-3333
;  INFORMATION FOR SEQ ID NO:  103:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  58 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-417-210A-103

  Query Match             0.9%;  Score 21.2;  DB 2;  Length 58;
  Best Local Similarity   64.0%;  Pred. No. 9.4e+03;
  Matches   32;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy   2221 CAACTTCTGTTTCTTCTCTTGGTGAAATAATAAAATGCAAATGAATCATT 2270
          | |  |||| ||||||| || |||  ||  || | || || |  ||  ||
Db     51 CTAGCTCTGCTTCTTCTGTTAGTGGCATTTTATATTGTAATTATATATTT 2


RESULT 36
US-08-417-210A-108
; Sequence 108, Application US/08417210A
; Patent No. 5863542
;  GENERAL INFORMATION:
;    APPLICANT:  PAOLETTI, ENZO
;    APPLICANT:  TARTAGLIA, JAMES
;    APPLICANT:  COX, WILLIAM I.
;    TITLE OF INVENTION:  IMMUNODEFICIENCY RECOMBINANT POXVIRUS
;    NUMBER OF SEQUENCES:  148
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CURTIS, MORRIS & SAFFORD, P.C.
;      STREET:  530 FIFTH AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/417,210A
;      FILING DATE:  05-APR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
```

```
;       NAME:  KOWALSKI, THOMAS J.
;       REGISTRATION NUMBER:  32,147
;       REFERENCE/DOCKET NUMBER:  454310-2690
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-840-3333
;  INFORMATION FOR SEQ ID NO:  108:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  58 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-417-210A-108

  Query Match             0.9%;  Score 21.2;  DB 2;  Length 58;
  Best Local Similarity   64.0%;  Pred. No. 9.4e+03;
  Matches   32;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy   2221 CAACTTCTGTTTCTTCTCTTGGTGAAATAATAAAATGCAAATGAATCATT 2270
          | |  |||| ||||||| || |||  ||  || | || || |  ||  ||
Db      7 CTAGCTCTGCTTCTTCTGTTAGTGGCATTTTATATTGTAATTATATATTT 56


RESULT 37
US-09-532-656-17
; Sequence 17, Application US/09532656
; Patent No. 6316608
; GENERAL INFORMATION:
;  APPLICANT: Reynolds, Mark A.
;  APPLICANT:  Ruvolo, Michael
;  APPLICANT:  Arnold, Jr., Lyle J.
;  TITLE OF INVENTION: COMBINED POLYNUCLEOTIDE SEQUENCES AS DISCRETE ASSAY ENDPOINTS
;  FILE REFERENCE: IN-0017 US
;  CURRENT APPLICATION NUMBER: US/09/532,656
;  CURRENT FILING DATE:  2000-03-20
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PERL Program
; SEQ ID NO 17
;   LENGTH: 59
;   TYPE: DNA
;   ORGANISM: Candida albicans
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6316608 F23S805
US-09-532-656-17

  Query Match             0.9%;  Score 21;  DB 4;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+04;
  Matches   30;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   1078 TTCTTTAGACACATTAACTGGGAAGAACTTCTGGCTCGAAAGGTG 1122
          ||| ||||  |||||||  |   | ||||  ||  | |||| |||
Db      5 TTCATTAGTGACATTAAAAGAAGAAAACTGATGTTTTGAAATGTG 49


RESULT 38
US-09-775-319-4
; Sequence 4, Application US/09775319
; Patent No. 6387631
; GENERAL INFORMATION:
;  APPLICANT: Arnold,Jr., Lyle J.
;  APPLICANT:  Sawan,Samuel P.
;  APPLICANT:  Lee, Paul H.
;  TITLE OF INVENTION: POLYMER COATED SURFACES FOR MICROARRAY APPLICATIONS
;  FILE REFERENCE: IN-0038 US
;  CURRENT APPLICATION NUMBER: US/09/775,319
;  CURRENT FILING DATE:  2001-02-01
;  PRIOR APPLICATION NUMBER: 09/532,419
;  PRIOR FILING DATE: 2000-12-20
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PERL Program
; SEQ ID NO 4
;   LENGTH: 59
;   TYPE: DNA
;   ORGANISM: Candida albicans
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6387631 F23S805
US-09-775-319-4

  Query Match             0.9%;  Score 21;  DB 4;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+04;
  Matches   30;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   1078 TTCTTTAGACACATTAACTGGGAAGAACTTCTGGCTCGAAAGGTG 1122
          ||| ||||  |||||||  |   | ||||  ||  | |||| |||
Db      5 TTCATTAGTGACATTAAAAGAAGAAAACTGATGTTTTGAAATGTG 49


RESULT 39
US-09-532-419A-4
; Sequence 4, Application US/09532419A
; Patent No. 6413722
; GENERAL INFORMATION:
;  APPLICANT:  Arnold,Jr., Lyle J.
;  APPLICANT:  Sawan,Samuel P.
;  APPLICANT:  Lee, Paul H.
;  TITLE OF INVENTION:  POLYMER COATED SURFACES FOR MICROARRAY APPLICATIONS
;  FILE REFERENCE:  IN-0038 US
;  CURRENT APPLICATION NUMBER: US/09/532,419A
;  CURRENT FILING DATE:  2000-03-22
;  NUMBER OF SEQ ID NOS:  10
;  SOFTWARE:  PERL Program
; SEQ ID NO 4
;   LENGTH: 59
;   TYPE: DNA
;   ORGANISM: Candida albicans
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6413722 F23S805
US-09-532-419A-4

  Query Match             0.9%;  Score 21;  DB 4;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+04;
  Matches   30;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   1078 TTCTTTAGACACATTAACTGGGAAGAACTTCTGGCTCGAAAGGTG 1122
          ||| ||||  |||||||  |   | ||||  ||  | |||| |||
Db      5 TTCATTAGTGACATTAAAAGAAGAAAACTGATGTTTTGAAATGTG 49


RESULT 40
US-08-475-081-2
; Sequence 2, Application US/08475081
; Patent No. 5948894
;  GENERAL INFORMATION:
;    APPLICANT:  Berry, Mark J.
;    APPLICANT:  Davis, Paul J.
;    APPLICANT:  Verhoeyen, Martine E.
;    APPLICANT:  De Winter, Ronald F.J.
;    TITLE OF INVENTION:  REAGENTS
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN, DARBY & CUSHMAN
;      STREET:  Eleventh Floor, 1615 L. Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20036-5601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,081
```

```
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/979,448
;       FILING DATE:  20-NOV-1992
;       APPLICATION NUMBER:  US/07/628,910
;       FILING DATE:  18-DEC-1990
;       APPLICATION NUMBER:  GB 8928501.9
;       FILING DATE:  18-DEC-1990
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 861-3000
;       TELEFAX:  (202) 822-0944
;       TELEX:  248453 CUSH
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  60 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     FEATURE:
;       NAME/KEY:  misc_difference
;       LOCATION:  replace(37..42, "")
;       OTHER INFORMATION:  /note= "nnn nnn represents (GCT
;       OTHER INFORMATION:  ACC)n where n=0 to 5"
US-08-475-081-2

  Query Match             0.9%;  Score 21;  DB 2;  Length 60;
  Best Local Similarity   62.8%;  Pred. No. 1.1e+04;
  Matches   27;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      58 TACCCAGCCCCGGACTTCCGAGACAGGGAAGCTGAGGACATGG 100
           |||||    ||||| ||||        |||  ||||||| | ||
Db      17 TACCCCTTACCGGAATTCCCNNNNNNGGATCCTGAGGAGACGG 59


RESULT 41
US-09-641-638-1114
; Sequence 1114, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Cohen, Annick
;  TITLE OF INVENTION:  BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
;  TITLE OF INVENTION:  GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
;  FILE REFERENCE:  GENSET.051CP1
;  CURRENT APPLICATION NUMBER: US/09/641,638
;  CURRENT FILING DATE:  2000-08-16
;  PRIOR APPLICATION NUMBER:  US 09/502,330
;  PRIOR FILING DATE:  2000-02-11
;  PRIOR APPLICATION NUMBER:  US 60/133,200
;  PRIOR FILING DATE:  1999-05-07
;  PRIOR APPLICATION NUMBER:  US 09/275,267
;  PRIOR FILING DATE:  1999-03-23
;  PRIOR APPLICATION NUMBER:  US 60/119,917
;  PRIOR FILING DATE:  1999-02-12
;  NUMBER OF SEQ ID NOS:  1304
;  SOFTWARE:  Patent.pm
; SEQ ID NO 1114
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 12-316-292 : polymorphic base C or T
;   NAME/KEY: misc_feature
;   LOCATION: 47
;   OTHER INFORMATION: n=a, g, c or t
US-09-641-638-1114

  Query Match             0.9%;  Score 20.6;  DB 4;  Length 47;
  Best Local Similarity   67.4%;  Pred. No. 1.2e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy     850 GATTGGTGGAGTTTGGGAGCATTAATGTATGACATGCTGACTG 892
           || |||||   | |||| || |  || | | ||||| || |||
Db       4 GAATGGTGTCCTCTGGGTGCYTGCATTTCTTACATGATGGCTG 46


RESULT 42
US-08-860-038-17
; Sequence 17, Application US/08860038
; Patent No. 6287762
;  GENERAL INFORMATION:
;    APPLICANT:  CROUZET, Joel
;    APPLICANT:  SCHERMAN, Daniel
;    APPLICANT:  WILS, Pierre
;    TITLE OF INVENTION:  PURIFICATION OF A TRIPLE HELIX FORMATION
;    TITLE OF INVENTION:  WITH AN IMMOBILIZED OLIGONUCLEOTIDE
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rhone-Poulenc Rorer Inc.
;      STREET:  500 Arcola Road, Mailstop 3C43
;      CITY:  Collegeville
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19426
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/860,038
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94/15162
;      FILING DATE:  16-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO FR95/01468
;      FILING DATE:  08-NOV-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Savitzky Esq., Martin F.
;      REGISTRATION NUMBER:  29,699
;      REFERENCE/DOCKET NUMBER:  ST94090-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610) 454-3816
;      TELEFAX:  (610) 454-3808
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "Oligonucleotide"
US-08-860-038-17

  Query Match             0.9%;  Score 20.6;  DB 4;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 1.3e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy     132 AGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAG 174
           |||||| | ||     ||||| |||  |||||||| ||   ||
Db       3 AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 45


RESULT 43
US-09-580-923-17
; Sequence 17, Application US/09580923
; Patent No. 6319672
; GENERAL INFORMATION:
;  APPLICANT:  Crouzet, Joel
```

```
;  APPLICANT:  Scherman, Daniel
;  APPLICANT:  Wils, Pierre
;  APPLICANT:  Cameron, Beatrice
;  APPLICANT:  Blanche, Francis
;  TITLE OF INVENTION:  PURIFICATION OF A TRIPLE HELIX FORMATION WITH AN
;  TITLE OF INVENTION:  IMMOBILIZED OLIGONUCLEOTIDE
;  FILE REFERENCE:  03804.0138-01
;  CURRENT APPLICATION NUMBER: US/09/580,923
;  CURRENT FILING DATE:  2000-05-26
;  PRIOR APPLICATION NUMBER:  08/860,038
;  PRIOR FILING DATE:  1997-06-09
;  PRIOR APPLICATION NUMBER:  PCT/FR95/01468
;  PRIOR FILING DATE:  1995-11-08
;  NUMBER OF SEQ ID NOS:  36
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 17
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: oligonucleotide
US-09-580-923-17

  Query Match             0.9%;  Score 20.6;  DB 4;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 1.3e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy     132 AGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAG 174
           |||||| | ||     ||||| |||  |||||||| ||   ||
Db       3 AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGAG 45


RESULT 44
US-08-929-501-24
; Sequence 24, Application US/08929501
; Patent No. 5854051
;  GENERAL INFORMATION:
;    APPLICANT:  Ramaswamy Chandrashekar
;    APPLICANT:  Naotoshi Tsuji,
;    TITLE OF INVENTION:  PARASITIC HELMINTH ASPARAGINASE
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/929,501
;      FILING DATE:  15-SEP-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  51 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Primer
US-08-929-501-24

  Query Match             0.9%;  Score 20.6;  DB 2;  Length 51;
  Best Local Similarity   67.4%;  Pred. No. 1.3e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    1848 CATCTTTCTCAACCTTATCAAGGATTTTCATGTTGATGACTCG 1890
           | |||| || ||| || |  |   |||||||  | |||||| |
Db       9 CTTCTTACTGAACTTTTTTCATCTTTTTCATTCTAATGACTAG 51


RESULT 45
US-09-140-177-24
; Sequence 24, Application US/09140177
; Patent No. 6042825
;  GENERAL INFORMATION:
;    APPLICANT:  Ramaswamy Chandrashekar
;    APPLICANT:  Naotoshi Tsuji,
;    TITLE OF INVENTION:  PARASITIC HELMINTH ASPARAGINASE
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/140,177
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/929,501
;      FILING DATE:  15-SEP-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  51 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Primer
US-09-140-177-24

  Query Match             0.9%;  Score 20.6;  DB 3;  Length 51;
  Best Local Similarity   67.4%;  Pred. No. 1.3e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    1848 CATCTTTCTCAACCTTATCAAGGATTTTCATGTTGATGACTCG 1890
           | |||| || ||| || |  |   |||||||  | |||||| |
Db       9 CTTCTTACTGAACTTTTTTCATCTTTTTCATTCTAATGACTAG 51


RESULT 46
US-09-397-979-24
; Sequence 24, Application US/09397979
; Patent No. 6165735
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Ramaswamy Chandrashekar
;    APPLICANT:  Naotoshi Tsuji,
;    TITLE OF INVENTION:  PARASITIC HELMINTH ASPARAGINASE
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/397,979
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/929,501
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  51 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Primer
US-09-397-979-24

  Query Match             0.9%;  Score 20.6;  DB 4;  Length 51;
  Best Local Similarity   67.4%;  Pred. No. 1.3e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1848 CATCTTTCTCAACCTTATCAAGGATTTTCATGTTGATGACTCG 1890
          | |||| || ||| || |  |   |||||||  | |||||| |
Db      9 CTTCTTACTGAACTTTTTTCATCTTTTTCATTCTAATGACTAG 51


RESULT 47
US-08-171-389-19
; Sequence 19, Application US/08171389
; Patent No. 5578444
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/171,389
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175/G19P3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human alcohal dehydrogenase beta
;      INDIVIDUAL ISOLATE:  subunit gene
US-08-171-389-19

  Query Match             0.9%;  Score 20.2;  DB 1;  Length 48;
  Best Local Similarity   68.3%;  Pred. No. 1.6e+04;
  Matches   28;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   2082 GATGTTTTACGTGCAAACAACCTGAATCTTTTTTTTATATA 2122
          |||||| ||  ||||||||  | ||| | | |   |||||
Db      1 GATGTTACACAAGCAAACAAAATAAATATCTGTGCAATATA 41


RESULT 48
US-08-123-936-19
; Sequence 19, Application US/08123936
; Patent No. 5726014
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    TITLE OF INVENTION:  Screening Assay for the Detection of
;    TITLE OF INVENTION:  DNA-Binding Molecules
;    NUMBER OF SEQUENCES:  640
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/123,936
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0075.32/G19P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human alcohal dehydrogenase beta
;      INDIVIDUAL ISOLATE:  subunit gene
US-08-123-936-19

  Query Match             0.9%;  Score 20.2;  DB 1;  Length 48;
  Best Local Similarity   68.3%;  Pred. No. 1.6e+04;
  Matches   28;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   2082 GATGTTTTACGTGCAAACAACCTGAATCTTTTTTTTATATA 2122
          ||||||  ||  ||||||||  | ||| | |  |  |||||
Db      1 GATGTTACACAAGCAAACAAAATAAATATCTGTGCAATATA 41


RESULT 49
US-08-475-228A-19
; Sequence 19, Application US/08475228A
; Patent No. 5869241
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,228A
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stratford, Carol A.
;      REGISTRATION NUMBER:  34,444
;      REFERENCE/DOCKET NUMBER:  4600-0175.21/G19P3D2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human alcohal dehydrogenase beta
;      INDIVIDUAL ISOLATE:  subunit gene
US-08-475-228A-19

  Query Match             0.9%;  Score 20.2;  DB 2;  Length 48;
  Best Local Similarity   68.3%;  Pred. No. 1.6e+04;
  Matches   28;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   2082 GATGTTTTACGTGCAAACAACCTGAATCTTTTTTTTATATA 2122
          ||||||  ||  ||||||||  | ||| | |  |  |||||
Db      1 GATGTTACACAAGCAAACAAAATAAATATCTGTGCAATATA 41


RESULT 50
US-08-482-080A-19
; Sequence 19, Application US/08482080A
; Patent No. 6010849
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,080A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/171,389
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  US 07/723,618
;       FILING DATE:  27-JUN-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/081,070
;       FILING DATE:  22-JUN-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Brady, John F.
;       REGISTRATION NUMBER:  39,118
;       REFERENCE/DOCKET NUMBER:  4600-0175.20/G19P3D1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (650) 324-0880
;       TELEFAX:  (650) 324-0960
;   INFORMATION FOR SEQ ID NO:  19:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  48 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ORIGINAL SOURCE:
;       INDIVIDUAL ISOLATE:  Human alcohal dehydrogenase beta
;       INDIVIDUAL ISOLATE:  subunit gene
US-08-482-080A-19

  Query Match             0.9%;  Score 20.2;  DB 3;  Length 48;
  Best Local Similarity   68.3%;  Pred. No. 1.6e+04;
  Matches   28;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    2082 GATGTTTTACGTGCAAACAACCTGAATCTTTTTTTTATATA 2122
           ||||||  ||  ||||||||  | ||| | | |   |||||
Db       1 GATGTTACACAAGCAAACAAAATAAATATCTGTGCAATATA 41


RESULT 51
US-09-354-947-19
; Sequence 19, Application US/09354947
; Patent No. 6384208
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/354,947
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/482,080
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/171,389
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brady, John F.
;      REGISTRATION NUMBER:  39,118
;      REFERENCE/DOCKET NUMBER:  4600-0175.20/G19P3D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 324-0880
;      TELEFAX:  (650) 324-0960
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human alcohal dehydrogenase beta
;      INDIVIDUAL ISOLATE:  subunit gene
US-09-354-947-19

  Query Match             0.9%;  Score 20.2;  DB 4;  Length 48;
  Best Local Similarity   68.3%;  Pred. No. 1.6e+04;
  Matches   28;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    2082 GATGTTTTACGTGCAAACAACCTGAATCTTTTTTTTATATA 2122
           ||||||  ||  ||||||||  | ||| | | |   |||||
Db       1 GATGTTACACAAGCAAACAAAATAAATATCTGTGCAATATA 41


RESULT 52
PCT-US93-12388-19
; Sequence 19, Application PC/TUS9312388
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12388
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175.41/G19PCT2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  19:
```

```
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  48 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ORIGINAL SOURCE:
;       INDIVIDUAL ISOLATE:  Human alcohal dehydrogenase beta
;       INDIVIDUAL ISOLATE:  subunit gene
PCT-US93-12388-19

  Query Match             0.9%;  Score 20.2;  DB 5;  Length 48;
  Best Local Similarity   68.3%;  Pred. No. 1.6e+04;
  Matches   28;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   2082 GATGTTTTACGTGCAAACAACCTGAATCTTTTTTTTATATA 2122
          ||||||  ||  ||||||||  | ||| | | |   |||||
Db      1 GATGTTACACAAGCAAACAAAATAAATATCTGTGCAATATA 41


RESULT 53
US-08-171-382-11/c
; Sequence 11, Application US/08171382
; Patent No. 5472856
;  GENERAL INFORMATION:
;    APPLICANT:  Harris, Crafford A.
;    APPLICANT:  Goldstein, Gideon
;    APPLICANT:  Siekierka, John J.
;    APPLICANT:  Talle, Mary Anne
;    APPLICANT:  Shenbagamurthi, Ponniah
;    APPLICANT:  Culler, Michael D.
;    APPLICANT:  Setcavage, Diane R.
;    TITLE OF INVENTION:  Recombinant Human Thymopoietin Proteins
;    TITLE OF INVENTION:  and Uses Therefor
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, P.O. Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/171,382
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  IRI43USA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9206
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
US-08-171-382-11

  Query Match             0.9%;  Score 20.2;  DB 1;  Length 50;
  Best Local Similarity   68.3%;  Pred. No. 1.6e+04;
  Matches   28;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   2226 TCTGTTTCTTCTCTTGGTGAAATAATAAAATGCAAATGAAT 2266
          ||| |||| || || |||    | ||||| ||| |||||
Db     42 TCTAATTCTCTTCATGTTGACCCTAGAAAATCCAACTGAAT 2


RESULT 54
US-08-584-760A-33/c
; Sequence 33, Application US/08584760A
; Patent No. 6290953
;  GENERAL INFORMATION:
;    APPLICANT:  Ballance, David J
;    APPLICANT:  Courtney, Michael G
;    APPLICANT:  Finnis, Christopher J A
;    APPLICANT:  Sleep, Darrell
;    TITLE OF INVENTION:  Medicine
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Centeon L.L.C.
;      STREET:  1020 First Avenue
;      CITY:  King of Prussia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19406-1310
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/584,760A
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/211,860
;      FILING DATE:  15-APR-1994
;      APPLICATION NUMBER:  GB 9121815.6
;      FILING DATE:  14-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Naomi Biswas
;      REGISTRATION NUMBER:  38,384
;      REFERENCE/DOCKET NUMBER:  92H853-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610/878-4294
;      TELEFAX:  610/878/4221
;  INFORMATION FOR SEQ ID NO:  33:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  60 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..60
;      OTHER INFORMATION:  /function= "oligonucleotide 32"
US-08-584-760A-33

  Query Match             0.9%;  Score 20.2;  DB 4;  Length 60;
  Best Local Similarity   63.3%;  Pred. No. 1.8e+04;
  Matches   31;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy    378 GAAAATATTTGCCATGAAGGTGCTTAAAAAGGCAATGATAGTAAGAAAT 426
          ||||| |  || || |||| |||| ||    |  || | |  ||  ||||
Db     58 GAAAACAACTGGCAAGAAGCTGCTAAAGCTGTTAAAGCTGCTATTAAAT 10


RESULT 55
US-08-222-177A-349
; Sequence 349, Application US/08222177A
; Patent No. 5582979
;  GENERAL INFORMATION:
;    APPLICANT:  Weber, James L.
```

```
;    TITLE OF INVENTION:  LENGTH POLYMORPHISMS IN
;    TITLE OF INVENTION:  (dC-dA)n.(dG-dT)n SEQUENCES AND METHODS OF USING SAME
;    NUMBER OF SEQUENCES:  460
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DeWitt Ross & Stevens, S.C.
;      STREET:  8000 Excelsior Drive, Suite 401
;      CITY:  Madison
;      STATE:  Wisconsin
;      COUNTRY:  USA
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/222,177A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/341,562
;      FILING DATE:  21-APR-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sara, Charles S.
;      REGISTRATION NUMBER:  30,492
;      REFERENCE/DOCKET NUMBER:  09865.601
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  349:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    IMMEDIATE SOURCE:
;      CLONE:  mfd110rs
US-08-222-177A-349

  Query Match             0.9%;  Score 20;  DB 1;  Length 46;
  Best Local Similarity   65.9%;  Pred. No. 1.8e+04;
  Matches   29;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2111 TTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCA 2154
          ||| ||||| ||  ||| ||||| |  | |   || || ||| |
Db      2 TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 45


RESULT 56
US-09-338-907-315
; Sequence 315, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT: Blumenfeld, Marta
;  APPLICANT: Ilya, Chumakov
;  APPLICANT: Bougueleret, Lydie
;  TITLE OF INVENTION: PROSTATE CANCER GENE
;  FILE REFERENCE: GENSET.18CP1CP
;  CURRENT APPLICATION NUMBER: US/09/338,907
;  CURRENT FILING DATE: 1999-06-23
;  EARLIER APPLICATION NUMBER: 08/996,306
;  EARLIER FILING DATE: 1997-12-22
;  EARLIER APPLICATION NUMBER: 60/099,658
;  EARLIER FILING DATE: 1998-09-09
;  EARLIER APPLICATION NUMBER: 09/218,207
;  EARLIER FILING DATE: 1998-12-22
;  NUMBER OF SEQ ID NOS: 578
;  SOFTWARE: Patent.pm
; SEQ ID NO 315
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 1..47
;   OTHER INFORMATION: polymorphic fragment 99-147-181, variant version of SEQ ID238
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: base G ; A in SEQ ID238
;   FEATURE:
;   NAME/KEY: primer_bind
;   LOCATION: 1..23
;   OTHER INFORMATION: potential microsequencing oligo 99-147-181.mis1
;   FEATURE:
;   NAME/KEY: primer_bind
;   LOCATION: 25..47
;   OTHER INFORMATION: complement potential microsequencing oligo 99-147-181.mis2
US-09-338-907-315

  Query Match             0.9%;  Score 20;  DB 4;  Length 47;
  Best Local Similarity   72.2%;  Pred. No. 1.8e+04;
  Matches   26;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy    397 GTGCTTAAAAAGGCAATGATAGTAAGAAATGCTAAA 432
          ||  | ||||||   ||||||| ||||||  || ||
Db      3 GTCATGAAAAAGAGCATGATAGAAAGAAAAACTTAA 38


RESULT 57
US-09-218-207-315
; Sequence 315, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT: Blumenfeld, Marta
;  APPLICANT: Ilya, Chumakov
;  APPLICANT: Bougueleret, Lydie
;  TITLE OF INVENTION: Prostate cancer gene
;  FILE REFERENCE: GENSET.018CP1
;  CURRENT APPLICATION NUMBER: US/09/218,207
;  CURRENT FILING DATE: 1998-12-22
;  EARLIER APPLICATION NUMBER: 08/996,306
;  EARLIER FILING DATE: 1997-12-22
;  EARLIER APPLICATION NUMBER: 60/099,658
;  EARLIER FILING DATE: 1998-09-09
;  NUMBER OF SEQ ID NOS: 578
;  SOFTWARE: Patent.pm
; SEQ ID NO 315
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 1..47
;   OTHER INFORMATION: polymorphic fragment 99-147-181, variant version of SEQ ID238
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: base G ; A in SEQ ID238
;   FEATURE:
;   NAME/KEY: primer_bind
;   LOCATION: 1..23
;   OTHER INFORMATION: potential microsequencing oligo 99-147-181.mis1
;   FEATURE:
;   NAME/KEY: primer_bind
;   LOCATION: 25..47
;   OTHER INFORMATION: complement potential microsequencing oligo 99-147-181.mis2
US-09-218-207-315

  Query Match             0.9%;  Score 20;  DB 4;  Length 47;
  Best Local Similarity   72.2%;  Pred. No. 1.8e+04;
  Matches   26;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;
```

```
Qy    397 GTGCTTAAAAAGGCAATGATAGTAAGAAATGCTAAA 432
          ||  | ||||||   ||||||| ||||||  || ||
Db      3 GTCATGAAAAAGAGCATGATAGAAAGAAAAACTTAA 38


RESULT 58
US-08-105-483-102/c
; Sequence 102, Application US/08105483
; Patent No. 5494807
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  462
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/105,483
;      FILING DATE:  12-AUG-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,951
;      FILING DATE:  06-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2400
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  102:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  49 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-105-483-102

  Query Match             0.9%;  Score 20;  DB 1;  Length 49;
  Best Local Similarity   65.9%;  Pred. No. 1.8e+04;
  Matches   29;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2109 CTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
          | |||| |  | ||| ||||||||||  |  | ||  | ||| |
Db     45 CATTTTATATTGTAATTATATATTTTCAATTTTGAGATCTGAAT 2


RESULT 59
US-08-709-209-102/c
; Sequence 102, Application US/08709209
; Patent No. 5762938
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  462
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/709,209
;      FILING DATE:  21-AUG-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/105,483
;      FILING DATE:  12-AUG-1993
;      APPLICATION NUMBER:  US 07/847,951
;      FILING DATE:  06-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2400
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  102:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  49 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-709-209-102

  Query Match             0.9%;  Score 20;  DB 1;  Length 49;
  Best Local Similarity   65.9%;  Pred. No. 1.8e+04;
  Matches   29;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2109 CTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
          | |||| |  | ||| ||||||||||  |  | ||  | ||| |
Db     45 CATTTTATATTGTAATTATATATTTTCAATTTTGAGATCTGAAT 2


RESULT 60
US-08-458-101-102/c
; Sequence 102, Application US/08458101
; Patent No. 5766599
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Perkus, Marion E.
;    APPLICANT:  Taylor, Jill
;    APPLICANT:  Tartaglia, James
;    APPLICANT:  No. 5766599ton, Elizabeth K.
;    APPLICANT:  Riviere, Michel
;    APPLICANT:  de Taisne, Charles
;    APPLICANT:  Limbach, Keith J.
;    APPLICANT:  Johnson, Gerard P.
;    APPLICANT:  Pincus, Steven E.
;    APPLICANT:  Cox, William I.
;    APPLICANT:  Audonnet, Jean-Christophe Francis
;    APPLICANT:  Gettig, Russell Robert
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  467
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/458,101
;       FILING DATE:  01-JUN-1995
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Frommer, William S.
;       REGISTRATION NUMBER:  25,506
;       REFERENCE/DOCKET NUMBER:  454310-2740
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 840-3333
;       TELEFAX:  (212) 840-0712
;   INFORMATION FOR SEQ ID NO:  102:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  49 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-458-101-102

  Query Match             0.9%;  Score 20;  DB 1;  Length 49;
  Best Local Similarity   65.9%;  Pred. No. 1.8e+04;
  Matches   29;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2109 CTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
          | |||| |  | ||| ||||||||||  |  | ||  | ||| |
Db     45 CATTTTATATTGTAATTATATATTTTCAATTTTGAGATCTGAAT 2


RESULT 61
US-09-457-959-8
; Sequence 8, Application US/09457959
; Patent No. 6362002
; GENERAL INFORMATION:
;  APPLICANT: Denison, Timothy J. et al.
;  TITLE OF INVENTION: Characterization of Individual Polymer
;  TITLE OF INVENTION:  Molecules Based on Monomer-Interface Interactions
;  FILE REFERENCE: 00246/522001
;  CURRENT APPLICATION NUMBER: US/09/457,959
;  CURRENT FILING DATE:  1999-12-09
;  PRIOR APPLICATION NUMBER: US 09/098,142
;  PRIOR FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: US 08/405,735
;  PRIOR FILING DATE: 1995-03-17
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
US-09-457-959-8

  Query Match             0.9%;  Score 20;  DB 4;  Length 60;
  Best Local Similarity   58.3%;  Pred. No. 2e+04;
  Matches   35;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy   1741 AAAAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAA 1800
          ||||| |||     |    ||||||  ||  ||      |||||||||   |||  ||||
Db      1 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 60


RESULT 62
US-08-750-128-8
; Sequence 8, Application US/08750128
; Patent No. 5916773
;  GENERAL INFORMATION:
;    APPLICANT:  MELE Antonio,
;    APPLICANT:  DE SANTIS Rita,
;    APPLICANT:  CARLONI Cristina,
;    APPLICANT:  COSCARELLA Annamaria
;    TITLE OF INVENTION:  Hybrid molecule of formula  GM-CSF-L-EPO  or
;    TITLE OF INVENTION:  EPO-L-GM-CSF for hematopoietic stimulation
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LOWE, PRICE, LeBLANC & BECKER
;      STREET:  99 Canal Center Plaza, Suite 300
;      CITY:  Alexandria
;      STATE:  VIRGINIA
;      COUNTRY:  USA
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk, 3.50 inch.
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/750,128
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  IT FI 94 A 000106
;      FILING DATE:  27-MAY-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Robert L. Price
;      REGISTRATION NUMBER:  22,685
;      REFERENCE/DOCKET NUMBER:  2879-007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-684-1111
;      TELEFAX:  703-684-1124
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-08-750-128-8

  Query Match             0.8%;  Score 19.8;  DB 2;  Length 48;
  Best Local Similarity   63.8%;  Pred. No. 2.1e+04;
  Matches   30;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy    127 CAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCA 173
          | | | ||||| ||  |    |||||  |  | |||||||  || |||
Db      2 CCGGCGGAGGAGGCTCGGGAGGAGGAGGTTCCGGAGGAGCCGGCTCA 48


RESULT 63
US-09-461-697-233
; Sequence 233, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
;  APPLICANT: COGENT NEUROSCIENCE, Inc.
;  APPLICANT: Lo, Donald C.
;  APPLICANT: Barney, Shawn
;  APPLICANT: Thomas, Mary Beth
;  APPLICANT: Portbury, Stuart D.
;  APPLICANT: Puranam, Kasturi
;  APPLICANT: Katz, Lawrence C.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
;  TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
;  TITLE OF INVENTION: CELL DEATH
;  FILE REFERENCE: 10001-005-999
;  CURRENT APPLICATION NUMBER: US/09/461,697
;  CURRENT FILING DATE: 1999-12-14
;  NUMBER OF SEQ ID NOS: 466
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 233
;   LENGTH: 48
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-461-697-233
```

```
  Query Match             0.8%;  Score 19.8;  DB 4;  Length 48;
  Best Local Similarity   63.8%;  Pred. No. 2.1e+04;
  Matches   30;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   1642 TGGAGAGGGAGATGTGTGAGCATCCTGCAAGGTGAAACAAGACTCAA 1688
          || ||  ||| ||| |  || | |  | || | |||| |||| | ||
Db      2 TGAAGGTGGAAATGAGGAAGAAGCTGGAAAAGAGAAAGAAGATTTAA 48


RESULT 64
US-08-105-483-103
; Sequence 103, Application US/08105483
; Patent No. 5494807
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  462
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/105,483
;      FILING DATE:  12-AUG-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,951
;      FILING DATE:  06-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2400
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  103:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  49 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-105-483-103

  Query Match             0.8%;  Score 19.8;  DB 1;  Length 49;
  Best Local Similarity   63.8%;  Pred. No. 2.1e+04;
  Matches   30;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   2103 CTGAATCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTG 2149
          |||   | |||| |  | ||| ||||||||||  |  | ||  | ||
Db      3 CTGCCCCATTTTATATTGTAATTATATATTTTCAATTTTGAGATCTG 49


RESULT 65
US-08-709-209-103
; Sequence 103, Application US/08709209
; Patent No. 5762938
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  462
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/709,209
;      FILING DATE:  21-AUG-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/105,483
;      FILING DATE:  12-AUG-1993
;      APPLICATION NUMBER:  US 07/847,951
;      FILING DATE:  06-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2400
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  103:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  49 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-709-209-103

  Query Match             0.8%;  Score 19.8;  DB 1;  Length 49;
  Best Local Similarity   63.8%;  Pred. No. 2.1e+04;
  Matches   30;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   2103 CTGAATCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTG 2149
          |||   | |||| |  | ||| ||||||||||  |  | ||  | ||
Db      3 CTGCCCCATTTTATATTGTAATTATATATTTTCAATTTTGAGATCTG 49


RESULT 66
US-08-458-101-103
; Sequence 103, Application US/08458101
; Patent No. 5766599
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Perkus, Marion E.
;    APPLICANT:  Taylor, Jill
;    APPLICANT:  Tartaglia, James
;    APPLICANT:  No. 5766599ton, Elizabeth K.
;    APPLICANT:  Riviere, Michel
;    APPLICANT:  de Taisne, Charles
;    APPLICANT:  Limbach, Keith J.
;    APPLICANT:  Johnson, Gerard P.
;    APPLICANT:  Pincus, Steven E.
;    APPLICANT:  Cox, William I.
;    APPLICANT:  Audonnet, Jean-Christophe Francis
;    APPLICANT:  Gettig, Russell Robert
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  467
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
```

```
;       ZIP:  10036
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/458,101
;       FILING DATE:  01-JUN-1995
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Frommer, William S.
;       REGISTRATION NUMBER:  25,506
;       REFERENCE/DOCKET NUMBER:  454310-2740
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 840-3333
;       TELEFAX:  (212) 840-0712
;   INFORMATION FOR SEQ ID NO:  103:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  49 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-458-101-103

  Query Match             0.8%;  Score 19.8;  DB 1;  Length 49;
  Best Local Similarity   63.8%;  Pred. No. 2.1e+04;
  Matches   30;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   2103 CTGAATCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTG 2149
          |||   | |||| |  | ||| ||||||||||  |  | ||  | ||
Db      3 CTGCCCCATTTTATATTGTAATTATATATTTTCAATTTTGAGATCTG 49


RESULT 67
US-09-423-744A-13/c
; Sequence 13, Application US/09423744A
; Patent No. 6372500
;   GENERAL INFORMATION:
;        APPLICANT: HSC Research and Development Limited Partnership
;        TITLE OF INVENTION: Episomal Expression Cassettes for Gene
;                            Therapy
;        NUMBER OF SEQUENCES: 19
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Rockey, Milnamov & Katz, Ltd.
;             STREET: 180 N. Stetson Avenue, Suite 4700
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: USA
;             ZIP: 60601
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/423,744A
;             FILING DATE: 12-No. 6372500-1999
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: PCT/CA98/00478
;             FILING DATE: May 14, 1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Lisa V. Mueller
;             REFERENCE/DOCKET NUMBER: DWW6064P0020US
;   INFORMATION FOR SEQ ID NO: 13:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 54 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        FEATURE:
;             NAME/KEY:  -
;             LOCATION: 1..54
;             IDENTIFICATION METHOD:
;             OTHER INFORMATION: /note= "K183'SS synthetic DNA
;             oligo-nucleotide - amplification primer for PCR mutagenesis"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-423-744A-13

  Query Match             0.8%;  Score 19.8;  DB 4;  Length 54;
  Best Local Similarity   69.2%;  Pred. No. 2.2e+04;
  Matches   27;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   2097 AACAACCTGAATCTTTTTTTTATATAAATATATATTTTT 2135
          ||||||||   |||||||||| |  |  | | | ||| |
Db     48 AACAACCTCTCTCTTTTTTTTTTCAGGTCTGTTTTTAT 10


RESULT 68
US-09-415-784-11/c
; Sequence 11, Application US/09415784
; Patent No. 6391632
;   GENERAL INFORMATION:
;        APPLICANT: Dubensky Jr., Thomas W.
;                   Polo, John M.
;                   Belli, Barbara A.
;                   Schlesinger, Sondra
;                   Dryga, Sergey A.
;                   Frolov, Ilya
;        TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS-BASED VECTORS
;                            WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
;                            SYNTHESIS
;        NUMBER OF SEQUENCES: 125
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Seed Intellectual Property Law Group PLLC
;             STREET: 701 Fifth Avenue, Suite 6300
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/415,784
;             FILING DATE: 08-Oct-1999
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McMasters, David D.
;             REGISTRATION NUMBER: 33,963
;             REFERENCE/DOCKET NUMBER: 930049.457D1 /1196.006
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 11:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 58 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-415-784-11

  Query Match             0.8%;  Score 19.8;  DB 4;  Length 58;
  Best Local Similarity   63.8%;  Pred. No. 2.3e+04;
  Matches   30;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   1743 AAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAAC 1789
          ||  |||  | | |||||| |  ||||  ||      ||||||||||
Db     57 AACAAAATTTTGTTTTTAACATTTCAAAAAAAAAAAAAAAAAAAAC 11


RESULT 69
```

```
US-09-415-785A-11/c
; Sequence 11, Application US/09415785A
; Patent No. 6426196
;   GENERAL INFORMATION:
;        APPLICANT: Dubensky Jr., Thomas W.
;                   Polo, John M.
;                   Belli, Barbara A.
;                   Schlesinger, Sondra
;                   Dryga, Sergey A.
;                   Frolov, Ilya
;        TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS-BASED VECTORS
;                            WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
;                            SYNTHESIS
;        NUMBER OF SEQUENCES: 125
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Seed Intellectual Property Law Group PLLC
;             STREET: 701 Fifth Avenue, Suite 6300
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/415,785A
;             FILING DATE: 08-Oct-1999
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McMasters, David D.
;             REGISTRATION NUMBER: 33,963
;             REFERENCE/DOCKET NUMBER: 930049.457D1 /1196.006
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 11:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 58 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-415-785A-11

  Query Match             0.8%;  Score 19.8;  DB 4;  Length 58;
  Best Local Similarity   63.8%;  Pred. No. 2.3e+04;
  Matches   30;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   1743 AAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAAC 1789
          ||  |||  | | |||||| |  ||||  ||       ||||||||||
Db     57 AACAAAATTTTGTTTTTAACATTTCAAAAAAAAAAAAAAAAAAAAAC 11


RESULT 70
US-08-944-465-11/c
; Sequence 11, Application US/08944465
; Patent No. 6451592
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr., Thomas W.
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Belli, Barbara A.
;    APPLICANT:  Schlesinger, Sondra
;    APPLICANT:  Dryga, Sergey A.
;    APPLICANT:  Frolov, Ilya
;    TITLE OF INVENTION:  RECOMBINANT ALPHAVIRUS-BASED VECTORS
;    TITLE OF INVENTION:  WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
;    TITLE OF INVENTION:  SYNTHESIS
;    NUMBER OF SEQUENCES:  125
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed Intellectual Property Law Group PLLC
;      STREET:  701 Fifth Avenue, Suite 6300
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/944,465
;      FILING DATE:  06-Oct-1997
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.457C4 / 1196.005
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  58 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-944-465-11

  Query Match             0.8%;  Score 19.8;  DB 4;  Length 58;
  Best Local Similarity   63.8%;  Pred. No. 2.3e+04;
  Matches   30;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   1743 AAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAAC 1789
          ||  |||  | | |||||| |  ||||  ||       ||||||||||
Db     57 AACAAAATTTTGTTTTTAACATTTCAAAAAAAAAAAAAAAAAAAAAC 11


RESULT 71
US-09-415-868-11/c
; Sequence 11, Application US/09415868
; Patent No. 6458560
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr., Thomas W.
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Belli, Barbara A.
;    APPLICANT:  Schlesinger, Sondra
;    APPLICANT:  Dryga, Sergey A.
;    APPLICANT:  Frolov, Ilya
;    TITLE OF INVENTION:  RECOMBINANT ALPHAVIRUS-BASED VECTORS
;    TITLE OF INVENTION:  WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
;    TITLE OF INVENTION:  SYNTHESIS
;    NUMBER OF SEQUENCES:  125
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed Intellectual Property Law Group PLLC
;      STREET:  701 Fifth Avenue, Suite 6300
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/415,868
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/944,465
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
```

```
;       REGISTRATION NUMBER:  33,963
;       REFERENCE/DOCKET NUMBER:  930049.457C4 / 1196.005
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;   INFORMATION FOR SEQ ID NO:  11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  58 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-09-415-868-11

  Query Match             0.8%;  Score 19.8;  DB 4;  Length 58;
  Best Local Similarity   63.8%;  Pred. No. 2.3e+04;
  Matches   30;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   1743 AAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAAC 1789
          ||  |||  | | ||||||  |  ||||  ||      ||||||||||
Db     57 AACAAAATTTTGTTTTTAACATTTCAAAAAAAAAAAAAAAAAAAAAC 11


RESULT 72
US-09-415-900-11/c
; Sequence 11, Application US/09415900
; Patent No. 6465634.
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr., Thomas W.
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Belli, Barbara A.
;    APPLICANT:  Schlesinger, Sondra
;    APPLICANT:  Dryga, Sergey A.
;    APPLICANT:  Frolov, Ilya
;    TITLE OF INVENTION:  RECOMBINANT ALPHAVIRUS-BASED VECTORS
;    TITLE OF INVENTION:  WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
;    TITLE OF INVENTION:  SYNTHESIS
;    NUMBER OF SEQUENCES:  125
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed Intellectual Property Law Group PLLC
;      STREET:  701 Fifth Avenue, Suite 6300
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/415,900
;      FILING DATE:  08-Oct-1999
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.457D4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  58 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-415-900-11

  Query Match             0.8%;  Score 19.8;  DB 4;  Length 58;
  Best Local Similarity   63.8%;  Pred. No. 2.3e+04;
  Matches   30;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   1743 AAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAAC 1789
          ||  |||  | | ||||||  |  ||||  ||      ||||||||||
Db     57 AACAAAATTTTGTTTTTAACATTTCAAAAAAAAAAAAAAAAAAAAAC 11


RESULT 73
US-08-350-260A-250/c
; Sequence 250, Application US/08350260A
; Patent No. 5962255
;  GENERAL INFORMATION:
;    APPLICANT:  Winter, Gregory Paul
;    APPLICANT:  Griffiths, Andrew David
;    APPLICANT:  Williams, Samuel Cameron
;    APPLICANT:  Waterhouse, Peter
;    APPLICANT:  Nissim, Ahuva
;    APPLICANT:  Johnson, Kevin Stuart
;    APPLICANT:  Smith, Andrew John Hammond
;    TITLE OF INVENTION:  Methods for producing members of specific
;    TITLE OF INVENTION:  binding pairs
;    NUMBER OF SEQUENCES:  602
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David W. Clough
;      STREET:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/350,260A
;      FILING DATE:  05-DEC-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9110549.4
;      FILING DATE:  15-MAY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9206318.9
;      FILING DATE:  24-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB91/01134
;      FILING DATE:  10-JUL-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB92/00883
;      FILING DATE:  15-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB93/00605
;      FILING DATE:  24-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/150,002
;      FILING DATE:  31-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/307,619
;      FILING DATE:  16-SEP-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clough, David W
;      REGISTRATION NUMBER:  36,107
;      REFERENCE/DOCKET NUMBER:  28111/32372
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;  INFORMATION FOR SEQ ID NO:  250:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-350-260A-250

  Query Match             0.8%;  Score 19.6;  DB 2;  Length 48;
  Best Local Similarity   63.6%;  Pred. No. 2.3e+04;
```

```
  Matches   28;  Conservative    1;  Mismatches   15;  Indels    0;  Gaps    0;

Qy     435 TACAGCTCATACAAAAGCAGAACGGAATATTCTGGAGGAAGTAA 478
           | |||||:||| | |||||| |  | ||     |||||  |  ||
Db      45 TGCAGCTYATATACAAGCAGCAGCGTATTCGGCGGAGGGACCAA 2


RESULT 74
US-08-707-860D-3
; Sequence 3, Application US/08707860D
; Patent No. 6420629
; GENERAL INFORMATION:
;  APPLICANT:  Xue, Bao G
;  APPLICANT:  Newton, Craig H
;  APPLICANT:  Sutton, Benjamin C S
;  APPLICANT:  Gawley, John R
;  APPLICANT:  Ellis, David D
;  TITLE OF INVENTION:  A PROCESS OF INCREASING PLANT GROWTH AND YIELD AND
;  TITLE OF INVENTION:  MODIFYING CELLULOSE PRODUCTION IN PLANTS
;  FILE REFERENCE:  42748
;  CURRENT APPLICATION NUMBER: US/08/707,860D
;  CURRENT FILING DATE:  1996-09-09
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 52
;   TYPE: DNA
;   ORGANISM: Acetobacter xylinum
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (16)..(48)
US-08-707-860D-3

  Query Match             0.8%;  Score 19.6;  DB 4;  Length 52;
  Best Local Similarity   73.5%;  Pred. No. 2.4e+04;
  Matches   25;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy     377 GGAAAATATTTGCCATGAAGGTGCTTAAAAAGGC 410
           || ||||||| |  || |||   |||||||| ||
Db       5 GGTAAATATTAGTGATTAAGCCCCTTAAAAAAGC 38


RESULT 75
US-08-750-128-11
; Sequence 11, Application US/08750128
; Patent No. 5916773
;  GENERAL INFORMATION:
;    APPLICANT:  MELE Antonio,
;    APPLICANT:  DE SANTIS Rita,
;    APPLICANT:  CARLONI Cristina,
;    APPLICANT:  COSCARELLA Annamaria
;    TITLE OF INVENTION:  Hybrid molecule of formula  GM-CSF-L-EPO  or
;    TITLE OF INVENTION:  EPO-L-GM-CSF for hematopoietic stimulation
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LOWE, PRICE, LeBLANC & BECKER
;      STREET:  99 Canal Center Plaza, Suite 300
;      CITY:  Alexandria
;      STATE:  VIRGINIA
;      COUNTRY:  USA
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk, 3.50 inch.
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/750,128
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  IT FI 94 A 000106
;      FILING DATE:  27-MAY-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Robert L. Price
;      REGISTRATION NUMBER:  22,685
;      REFERENCE/DOCKET NUMBER:  2879-007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-684-1111
;      TELEFAX:  703-684-1124
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  57 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-08-750-128-11

  Query Match             0.8%;  Score 19.6;  DB 2;  Length 57;
  Best Local Similarity   66.7%;  Pred. No. 2.6e+04;
  Matches   28;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy     133 GAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAG 174
           ||||| ||  |   ||||| |  | |||||||  || ||||
Db       2 GAGGAGGCTCGGGAGGAGGAGGTTCCGGAGGAGCCGGCTCAG 43


RESULT 76
US-08-508-088-1
; Sequence 1, Application US/08508088
; Patent No. 5851826
;  GENERAL INFORMATION:
;    APPLICANT:  FRAEFEL, Cornel
;    APPLICANT:  GELLER, Alfred
;    APPLICANT:  LIM, Filip
;    TITLE OF INVENTION:  HELPER VIRUS-FREE HERPESVIRUS VECTOR PACKAGING SYSTEM
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DIKE, BRONSTEIN, ROBERTS & CUSHMAN
;      STREET:  130 WATER STREET
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02019
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/508,088
;      FILING DATE:  26-JUL-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DAVID, RESNICK S
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  45485
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:  200291 STRE
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  60 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
```

```
;    ORIGINAL SOURCE:
US-08-508-088-1

  Query Match             0.8%;  Score 19.6;  DB 2;  Length 60;
  Best Local Similarity   62.0%;  Pred. No. 2.6e+04;
  Matches   31;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       5 GAGGCTGCGGCGGGTCCGGGCCCATGAGGCGACGAAGGAGGCGGGACGGC 54
           | ||| |||| ||| | || | |  | | ||||   || || |||  | |
Db       2 GCGGCGGCGGTGGGCCGGGCCTCTGGCGCCGACTCGGGCGGGGGGCTGTC 51


RESULT 77
US-09-009-925-1
; Sequence 1, Application US/09009925
; Patent No. 5998208
;  GENERAL INFORMATION:
;    APPLICANT:  CHILDREN'S MEDICAL CENTER CORPORATION
;    TITLE OF INVENTION:  HELPER VIRUS-FREE HERPESVIRUS VECTOR PACKAGING SYSTEM
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
;      STREET:  130 WATER STREET
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02019
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/009,925
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/12033
;      FILING DATE:  22-JUL-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/508,088
;      FILING DATE:  26-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DAVID, RESNICK S
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  45485-CON
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  60 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
US-09-009-925-1

  Query Match             0.8%;  Score 19.6;  DB 2;  Length 60;
  Best Local Similarity   62.0%;  Pred. No. 2.6e+04;
  Matches   31;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       5 GAGGCTGCGGCGGGTCCGGGCCCATGAGGCGACGAAGGAGGCGGGACGGC 54
           | ||| |||| ||| | || | |  | | ||||   || || |||  | |
Db       2 GCGGCGGCGGTGGGCCGGGCCTCTGGCGCCGACTCGGGCGGGGGGCTGTC 51


RESULT 78
US-08-068-747-9/c
; Sequence 9, Application US/08068747
; Patent No. 5695933
;  GENERAL INFORMATION:
;    APPLICANT:  Schalling, Martin
;    APPLICANT:  Hudson, Thomas J.
;    APPLICANT:  Housman, David E.
;    TITLE OF INVENTION:  Direct Determination of Expanded
;    TITLE OF INVENTION:  Nucleotide Repeats in the Human Genome
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusett
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/068,747
;      FILING DATE:  28-MAY-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-6141
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  39 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "Synthetic"
US-08-068-747-9

  Query Match             0.8%;  Score 19.4;  DB 1;  Length 39;
  Best Local Similarity   70.3%;  Pred. No. 2.4e+04;
  Matches   26;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy     133 GAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGG 169
           ||||| | ||     ||||| |||  |||||||| ||
Db      37 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGAGG 1


RESULT 79
US-08-938-830-30
; Sequence 30, Application US/08938830
; Patent No. 6040437
;  GENERAL INFORMATION:
;    APPLICANT:  Lasky, Laurence A.
;    APPLICANT:  Dowbenko, Donald J.
;    TITLE OF INVENTION:  Tyrosine Phosphorylated Cleavage
;    TITLE OF INVENTION:  Furrow-Associated Proteins (PSTPIPs)
;    NUMBER OF SEQUENCES:  73
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  1 DNA Way
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
```

```
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/938,830
;       FILING DATE:
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/798419
;       FILING DATE:  07-FEB-1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Dreger, Ginger R.
;       REGISTRATION NUMBER:  33,055
;       REFERENCE/DOCKET NUMBER:  P1066P1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  650/225-3216
;       TELEFAX:  650/952-9881
;   INFORMATION FOR SEQ ID NO:  30:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  42 base pairs
;       TYPE:  Nucleic Acid
;       STRANDEDNESS:  Single
;       TOPOLOGY:  Linear
US-08-938-830-30

  Query Match             0.8%;  Score 19.4;  DB 3;  Length 42;
  Best Local Similarity   70.3%;  Pred. No. 2.5e+04;
  Matches   26;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      18 GTCCGGGCCCATGAGGCGACGAAGGAGGCGGGACGGC 54
           || |||  |||||| ||  |  |||  || |||||||
Db       3 GTTCGGATCCATGATGCTGCAGAGGCTGCTGGACGGC 39


RESULT 80
US-08-294-770A-6/c
; Sequence 6, Application US/08294770A
; Patent No. 5580754
;  GENERAL INFORMATION:
;    APPLICANT:  Babru, Samal B.
;    TITLE OF INVENTION:  Progenitor B Cell Stimulating Factor
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  1840 Dehavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/294,770A
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Winter, Robert B.
;      REFERENCE/DOCKET NUMBER:  A-214
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-294-770A-6

  Query Match             0.8%;  Score 19.4;  DB 1;  Length 45;
  Best Local Similarity   64.4%;  Pred. No. 2.6e+04;
  Matches   29;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    1377 ATTTTCTCCTGGGGATTTCTGGGGAAGAGGTGCTTCGGCCAGCAC 1421
           || |||||||||||  |    ||| | | |  ||  | | || |||
Db      45 ATCTTCTCCTGGGGGTACTGGGGCAGGGGAGGCAGCAGGCAACAC 1


RESULT 81
US-08-448-735C-6/c
; Sequence 6, Application US/08448735C
; Patent No. 5874399
;  GENERAL INFORMATION:
;    APPLICANT:  Babru, Samal B.
;    TITLE OF INVENTION:  Progenitor B Cell Stimulating Factor
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  1840 Dehavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/448,735C
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Winter, Robert B.
;      REFERENCE/DOCKET NUMBER:  A-214B
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-448-735C-6

  Query Match             0.8%;  Score 19.4;  DB 2;  Length 45;
  Best Local Similarity   64.4%;  Pred. No. 2.6e+04;
  Matches   29;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    1377 ATTTTCTCCTGGGGATTTCTGGGGAAGAGGTGCTTCGGCCAGCAC 1421
           || |||||||||||  |    ||| | | |  ||  | | || |||
Db      45 ATCTTCTCCTGGGGGTACTGGGGCAGGGGAGGCAGCAGGCAACAC 1


RESULT 82
US-09-641-638-1045
; Sequence 1045, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Cohen, Annick
;  TITLE OF INVENTION:  BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
;  TITLE OF INVENTION:  GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
;  FILE REFERENCE:  GENSET.051CP1
;  CURRENT APPLICATION NUMBER: US/09/641,638
;  CURRENT FILING DATE:  2000-08-16
;  PRIOR APPLICATION NUMBER:  US 09/502,330
;  PRIOR FILING DATE:  2000-02-11
;  PRIOR APPLICATION NUMBER:  US 60/133,200
;  PRIOR FILING DATE:  1999-05-07
;  PRIOR APPLICATION NUMBER:  US 09/275,267
;  PRIOR FILING DATE:  1999-03-23
;  PRIOR APPLICATION NUMBER:  US 60/119,917
;  PRIOR FILING DATE:  1999-02-12
;  NUMBER OF SEQ ID NOS:  1304
;  SOFTWARE:  Patent.pm
; SEQ ID NO 1045
```

```
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 12-888-98 : polymorphic base G or A
US-09-641-638-1045

  Query Match             0.8%;  Score 19.4;  DB 4;  Length 47;
  Best Local Similarity   74.2%;  Pred. No. 2.6e+04;
  Matches   23;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy   2120 ATAAATATATATTTTTCAAATAGATTTTTGA 2150
           ||||| || ||||||| ||| ||:||  | |
Db      1 ATAAAAATGTATTTTTGAAAAAGRTTCATAA 31


RESULT 83
US-08-983-607-5/c
; Sequence 5, Application US/08983607
; Patent No. 6140470
;  GENERAL INFORMATION:
;    APPLICANT:  Alan Garen
;    APPLICANT:  Xiaohong Cai
;    TITLE OF INVENTION:  Human Anti-Tumor Monoclonal Anti-
;    TITLE OF INVENTION:  bodies
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Department of Molecular Biophysics
;      ADDRESSEE:  and Biochemistry, Yale University
;      STREET:  266 Whitney Avenue
;      CITY:  New Haven
;      STATE:  Connecticut
;      COUNTRY:  United States of America
;      ZIP:  06520-8114
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" 1.44 Mb diskette
;      COMPUTER:  IBM PC
;      OPERATING SYSTEM:  MS DOS
;      SOFTWARE:  Word Processing
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/983,607
;      FILING DATE:  April 27, 1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/IB96/01032
;      FILING DATE:  June 28, 1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary M. Krinsky
;      REGISTRATION NUMBER:  32423
;      REFERENCE/DOCKET NUMBER:  OCR-679
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  203-773-9544
;      TELEFAX:  203-773-1183
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 residues
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;      DESCRIPTION:  primer used in constructs
US-08-983-607-5

  Query Match             0.8%;  Score 19.4;  DB 3;  Length 50;
  Best Local Similarity   60.5%;  Pred. No. 2.7e+04;
  Matches   26;  Conservative    3;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1501 GCATCGGCACCACTTCCAATACGACAGCCGAACTCTGGGCCAT 1543
          :|: | ||  :|| |     ||| | ||||||| ||| |||||
Db     43 SCWSCAGCTGYACCTGGGCCACGTCGGCCGAACCCTGAGCCAT 1


RESULT 84
US-08-983-607-42/c
; Sequence 42, Application US/08983607
; Patent No. 6140470
;  GENERAL INFORMATION:
;    APPLICANT:  Alan Garen
;    APPLICANT:  Xiaohong Cai
;    TITLE OF INVENTION:  Human Anti-Tumor Monoclonal Anti-
;    TITLE OF INVENTION:  bodies
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Department of Molecular Biophysics
;      ADDRESSEE:  and Biochemistry, Yale University
;      STREET:  266 Whitney Avenue
;      CITY:  New Haven
;      STATE:  Connecticut
;      COUNTRY:  United States of America
;      ZIP:  06520-8114
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" 1.44 Mb diskette
;      COMPUTER:  IBM PC
;      OPERATING SYSTEM:  MS DOS
;      SOFTWARE:  Word Processing
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/983,607
;      FILING DATE:  April 27, 1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/IB96/01032
;      FILING DATE:  June 28, 1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary M. Krinsky
;      REGISTRATION NUMBER:  32423
;      REFERENCE/DOCKET NUMBER:  OCR-679
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  203-773-9544
;      TELEFAX:  203-773-1183
;  INFORMATION FOR SEQ ID NO:  42:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 residues
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;      DESCRIPTION:  primer used in constructs
US-08-983-607-42

  Query Match             0.8%;  Score 19.4;  DB 3;  Length 50;
  Best Local Similarity   60.5%;  Pred. No. 2.7e+04;
  Matches   26;  Conservative    3;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1501 GCATCGGCACCACTTCCAATACGACAGCCGAACTCTGGGCCAT 1543
          :|: | ||  :|| |     ||| | ||||||| ||| |||||
Db     43 SCWSCAGCTGYACCTGGGCCACGTCGGCCGAACCCTGAGCCAT 1


RESULT 85
US-09-264-737-4
; Sequence 4, Application US/09264737A
; Patent No. 6107549
; GENERAL INFORMATION:
;  APPLICANT: Feng, Paul C.C.
;  APPLICANT: Ruff, Thomas G.
;  TITLE OF INVENTION: Engineering Plant Resistance to Pyridines via
;  TITLE OF INVENTION: Expression of Esterase Enzymes
;  FILE REFERENCE: 38-21(10551).RLE3 Pyridine Tolerance
;  CURRENT APPLICATION NUMBER: US/09/264,737A
;  CURRENT FILING DATE: 1999-03-09
;  EARLIER APPLICATION NUMBER: 60/077,377
;  EARLIER FILING DATE: 1998-03-10
```

```
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 55
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: amino terminal degenerate synthetic amplification
;   OTHER INFORMATION: primer based on published amino acid sequence of
;   OTHER INFORMATION: rabbit liver esterase isozyme 1 RLE1(Ozols, 1987);
;   OTHER INFORMATION: n44 & n50 = a, c, g, or t
US-09-264-737-4

  Query Match             0.8%;  Score 19.4;  DB 3;  Length 55;
  Best Local Similarity   65.7%;  Pred. No. 2.8e+04;
  Matches   23;  Conservative    2;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1270 GCTCCATCTGTACTTGAAAGTGTGAAAGAAAAGTT 1304
          || ||| ||||  |||| | ||| ||:|  ||: |
Db     21 GCACCACCTGTGGTTGACACTGTNAARGGNAARGT 55


RESULT 86
US-08-592-406-9/c
; Sequence 9, Application US/08592406
; Patent No. 5821059
;  GENERAL INFORMATION:
;    APPLICANT:  MINION, F. Chris
;    APPLICANT:  KNUDTSON, Kevin L.
;    TITLE OF INVENTION:  MYCOPLASMA EXPRESSION SYSTEM
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,406
;      FILING DATE:  06-FEB-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US93/07407
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  76645/132
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  60 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-592-406-9

  Query Match             0.8%;  Score 19.4;  DB 1;  Length 60;
  Best Local Similarity   60.4%;  Pred. No. 3e+04;
  Matches   32;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy   1799 AATAGTATTGCTGAACTCTTAGGCACATCAATTAATTGATTCCTCGCGACATC 1851
          | | ||||| || ||| |||  ||  ||  ||||  | ||   |    |||||
Db     59 ATTTGTATTCCTCAACACTTTCGCGTATATATTATATCATGTATGAATACATC 7


RESULT 87
US-08-741-881-44
; Sequence 44, Application US/08741881
; Patent No. 5789245
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr, Thomas W
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Ibanez, Carlos E.
;    APPLICANT:  Chang, Stephen M.W.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    APPLICANT:  Belli, Barbara A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/741,881
;      FILING DATE:  30-OCT-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.423C6 / 1146.007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  34 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-741-881-44

  Query Match             0.8%;  Score 19.2;  DB 1;  Length 34;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+04;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   2121 TAAATATATATTTTTCAAATAGATTTTTGATT 2152
          || ||||| || ||| | || |||| |||| |
Db      1 TATATATAGATCTTTGACATTGATTATTGACT 32


RESULT 88
US-08-739-158-44
; Sequence 44, Application US/08739158
; Patent No. 5814482
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr, Thomas W
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
```

```
;       ZIP:  98104-7092
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/739,158
;       FILING DATE:  30-OCT-1996
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  McMasters, David D.
;       REGISTRATION NUMBER:  33,963
;       REFERENCE/DOCKET NUMBER:  930049.423D3 / 1146.012
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;   INFORMATION FOR SEQ ID NO:  44:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  34 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-739-158-44

  Query Match             0.8%;  Score 19.2;  DB 1;  Length 34;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+04;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   2121 TAAATATATATTTTTCAAATAGATTTTTGATT 2152
          || ||||| || ||| | || |||| |||| |
Db      1 TATATATAGATCTTTGACATTGATTATTGACT 32


RESULT 89
US-08-739-167-44
; Sequence 44, Application US/08739167
; Patent No. 5843723
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr, Thomas W
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Ibanez, Carlos E.
;    APPLICANT:  Chang, Stephen M.W.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    APPLICANT:  Belli, Barbara A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS  AND ALPHAVIRUS
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/739,167
;      FILING DATE:  30-OCT-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.423C7 / 1146.008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH:  34 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-739-167-44

  Query Match             0.8%;  Score 19.2;  DB 2;  Length 34;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+04;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   2121 TAAATATATATTTTTCAAATAGATTTTTGATT 2152
          || ||||| || ||| | || |||| |||| |
Db      1 TATATATAGATCTTTGACATTGATTATTGACT 32


RESULT 90
US-08-404-796-44
; Sequence 44, Application US/08404796
; Patent No. 6015686
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr, Thomas W
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Ibanez, Carlos E.
;    APPLICANT:  Chang, Stephen M.W.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    APPLICANT:  Belli, Barbara A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/404,796
;      FILING DATE:  15-MAR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.423C5 / 1146.006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH:  34 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-404-796-44

  Query Match             0.8%;  Score 19.2;  DB 3;  Length 34;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+04;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   2121 TAAATATATATTTTTCAAATAGATTTTTGATT 2152
          || ||||| || ||| | || |||| |||| |
Db      1 TATATATAGATCTTTGACATTGATTATTGACT 32


RESULT 91
US-08-931-869-44
; Sequence 44, Application US/08931869
; Patent No. 6015694
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr, Thomas W
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Ibanez, Carlos E.
;    APPLICANT:  Chang, Stephen M.W.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    APPLICANT:  Belli, Barbara A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/931,869
;      FILING DATE:  16-SEP-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/404,796
;      FILING DATE:  15-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.423C5 / 1146.006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  34 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-931-869-44

  Query Match             0.8%;  Score 19.2;  DB 3;  Length 34;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+04;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   2121 TAAATATATATTTTTCAAATAGATTTTTGATT 2152
          || ||||| || ||| | || |||| |||| |
Db      1 TATATATAGATCTTTGACATTGATTATTGACT 32


RESULT 92
US-09-350-399-44
; Sequence 44, Application US/09350399
; Patent No. 6342372
;   GENERAL INFORMATION:
;        APPLICANT: Dubensky Jr, Thomas W
;                   Polo, John M.
;                   Jolly, Douglas J.
;                   Driver, David A.
;        TITLE OF INVENTION: EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;        NUMBER OF SEQUENCES: 128
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SEED and BERRY LLP
;             STREET: 6300 Columbia Center, 701 Fifth Avenue
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: US
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,399
;             FILING DATE: 08-Jul-1999
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McMasters, David D.
;             REGISTRATION NUMBER: 33,963
;             REFERENCE/DOCKET NUMBER: 930049.423D1 / 1146.010
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 44:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 34 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-09-350-399-44

  Query Match             0.8%;  Score 19.2;  DB 4;  Length 34;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+04;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   2121 TAAATATATATTTTTCAAATAGATTTTTGATT 2152
          || ||||| || ||| | || |||| |||| |
Db      1 TATATATAGATCTTTGACATTGATTATTGACT 32


RESULT 93
US-09-236-140A-44
; Sequence 44, Application US/09236140A
; Patent No. 6376236
;   GENERAL INFORMATION:
;        APPLICANT: Dubensky Jr, Thomas W
;                   Polo, John M.
;                   Ibanez, Carlos E.
;                   Chang, Stephen M.W.
;                   Jolly, Douglas J.
;                   Driver, David A.
;                   Belli, Barbara A.
;        TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS PARTICLES
;        NUMBER OF SEQUENCES: 124
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OPPENHEIMER WOLFF & DONNELLY
;             STREET: 840 NEWPORT CENTER DRIVE, SUITE 700
;             CITY: NEWPORT BEACH
;             STATE: CALIFORNIA
;             COUNTRY: US
;             ZIP: 92660
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/236,140A
;             FILING DATE: 22-Jan-1999
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cullman, Louis C.
;             REGISTRATION NUMBER: 39,645
;             REFERENCE/DOCKET NUMBER: 20263.332 / 1146.020
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (949) 823.6000
;             TELEFAX: (949) 823.6100
;   INFORMATION FOR SEQ ID NO: 44:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 34 base pairs
;             TYPE: nucleic acid
```

```
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-09-236-140A-44

  Query Match             0.8%;  Score 19.2;  DB 4;  Length 34;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+04;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   2121 TAAATATATATTTTTCAAATAGATTTTTGATT 2152
          || ||||| || ||| | || |||| |||| |
Db      1 TATATATAGATCTTTGACATTGATTATTGACT 32


RESULT 94
US-09-415-784-41
; Sequence 41, Application US/09415784
; Patent No. 6391632
;   GENERAL INFORMATION:
;        APPLICANT: Dubensky Jr., Thomas W.
;                   Polo, John M.
;                   Belli, Barbara A.
;                   Schlesinger, Sondra
;                   Dryga, Sergey A.
;                   Frolov, Ilya
;        TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS-BASED VECTORS
;                            WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
;                            SYNTHESIS
;        NUMBER OF SEQUENCES: 125
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Seed Intellectual Property Law Group PLLC
;             STREET: 701 Fifth Avenue, Suite 6300
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/415,784
;             FILING DATE: 08-Oct-1999
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McMasters, David D.
;             REGISTRATION NUMBER: 33,963
;             REFERENCE/DOCKET NUMBER: 930049.457D1 /1196.006
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 41:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 34 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 41:
US-09-415-784-41

  Query Match             0.8%;  Score 19.2;  DB 4;  Length 34;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+04;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   2121 TAAATATATATTTTTCAAATAGATTTTTGATT 2152
          || ||||| || ||| | || |||| |||| |
Db      1 TATATATAGATCTTTGACATTGATTATTGACT 32


RESULT 95
US-09-415-785A-41
; Sequence 41, Application US/09415785A
; Patent No. 6426196
;   GENERAL INFORMATION:
;        APPLICANT: Dubensky Jr., Thomas W.
;                   Polo, John M.
;                   Belli, Barbara A.
;                   Schlesinger, Sondra
;                   Dryga, Sergey A.
;                   Frolov, Ilya
;        TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS-BASED VECTORS
;                            WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
;                            SYNTHESIS
;        NUMBER OF SEQUENCES: 125
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Seed Intellectual Property Law Group PLLC
;             STREET: 701 Fifth Avenue, Suite 6300
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/415,785A
;             FILING DATE: 08-Oct-1999
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McMasters, David D.
;             REGISTRATION NUMBER: 33,963
;             REFERENCE/DOCKET NUMBER: 930049.457D1 /1196.006
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 41:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 34 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 41:
US-09-415-785A-41

  Query Match             0.8%;  Score 19.2;  DB 4;  Length 34;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+04;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   2121 TAAATATATATTTTTCAAATAGATTTTTGATT 2152
          || ||||| || ||| | || |||| |||| |
Db      1 TATATATAGATCTTTGACATTGATTATTGACT 32


RESULT 96
US-08-944-465-41
; Sequence 41, Application US/08944465
; Patent No. 6451592
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr., Thomas W.
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Belli, Barbara A.
;    APPLICANT:  Schlesinger, Sondra
;    APPLICANT:  Dryga, Sergey A.
;    APPLICANT:  Frolov, Ilya
;    TITLE OF INVENTION:  RECOMBINANT ALPHAVIRUS-BASED VECTORS
;    TITLE OF INVENTION:  WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
;    TITLE OF INVENTION:  SYNTHESIS
;    NUMBER OF SEQUENCES:  125
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed Intellectual Property Law Group PLLC
;      STREET:  701 Fifth Avenue, Suite 6300
;      CITY:  Seattle
;      STATE:  Washington
```

```
;       COUNTRY:  USA
;       ZIP:  98104-7092
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/944,465
;       FILING DATE:  06-Oct-1997
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  McMasters, David D.
;       REGISTRATION NUMBER:  33,963
;       REFERENCE/DOCKET NUMBER:  930049.457C4 / 1196.005
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;   INFORMATION FOR SEQ ID NO:  41:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  34 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-944-465-41

  Query Match             0.8%;  Score 19.2;  DB 4;  Length 34;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+04;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   2121 TAAATATATATTTTTCAAATAGATTTTTGATT 2152
          || ||||| || ||| | || |||| |||| |
Db      1 TATATATAGATCTTTGACATTGATTATTGACT 32


RESULT 97
US-09-415-868-41
; Sequence 41, Application US/09415868
; Patent No. 6458560
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr., Thomas W.
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Belli, Barbara A.
;    APPLICANT:  Schlesinger, Sondra
;    APPLICANT:  Dryga, Sergey A.
;    APPLICANT:  Frolov, Ilya
;    TITLE OF INVENTION:  RECOMBINANT ALPHAVIRUS-BASED VECTORS
;    TITLE OF INVENTION:  WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
;    TITLE OF INVENTION:  SYNTHESIS
;    NUMBER OF SEQUENCES:  125
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed Intellectual Property Law Group PLLC
;      STREET:  701 Fifth Avenue, Suite 6300
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/415,868
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/944,465
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.457C4 / 1196.005
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  41:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  34 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-415-868-41

  Query Match             0.8%;  Score 19.2;  DB 4;  Length 34;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+04;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   2121 TAAATATATATTTTTCAAATAGATTTTTGATT 2152
          || ||||| || ||| | || |||| |||| |
Db      1 TATATATAGATCTTTGACATTGATTATTGACT 32


RESULT 98
US-09-415-900-41
; Sequence 41, Application US/09415900
; Patent No. 6465634
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr., Thomas W.
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Belli, Barbara A.
;    APPLICANT:  Schlesinger, Sondra
;    APPLICANT:  Dryga, Sergey A.
;    APPLICANT:  Frolov, Ilya
;    TITLE OF INVENTION:  RECOMBINANT ALPHAVIRUS-BASED VECTORS
;    TITLE OF INVENTION:  WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
;    TITLE OF INVENTION:  SYNTHESIS
;    NUMBER OF SEQUENCES:  125
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed Intellectual Property Law Group PLLC
;      STREET:  701 Fifth Avenue, Suite 6300
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/415,900
;      FILING DATE:  08-Oct-1999
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.457D4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  41:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  34 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-415-900-41

  Query Match             0.8%;  Score 19.2;  DB 4;  Length 34;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+04;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   2121 TAAATATATATTTTTCAAATAGATTTTTGATT 2152
          || ||||| || ||| | || |||| |||| |
Db      1 TATATATAGATCTTTGACATTGATTATTGACT 32
```

```
RESULT 99
US-09-076-193-1
; Sequence 1, Application US/09076193
; Patent No. 5973231
;  GENERAL INFORMATION:
;    APPLICANT:  Bradfisch, Gregory A.
;    APPLICANT:  Muller-Cohn, Judy
;    APPLICANT:  Narva, Kenneth E.
;    APPLICANT:  Fu, Jenny M.
;    APPLICANT:  Thompson, Mark
;    TITLE OF INVENTION:  Bacillus thuringiensis Isolates, Toxins, and
;    TITLE OF INVENTION:  Genes for Controlling Certain Coleopteran Pests
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik, Lloyd & Saliwanchik
;      STREET:  2421 NW 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  FL
;      COUNTRY:  USA
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/076,193
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sanders, Jay M.
;      REGISTRATION NUMBER:  39,355
;      REFERENCE/DOCKET NUMBER:  MA-716
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (352) 375-8100
;      TELEFAX:  (352) 372-5800
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  40 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-076-193-1

  Query Match             0.8%;  Score 19.2;  DB 2;  Length 40;
  Best Local Similarity   60.0%;  Pred. No. 2.7e+04;
  Matches   24;  Conservative    4;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   2240 TGGTGAAATAATAAAATGCAAATGAATCATTGTTAACACA 2279
          |||  ||| ||| :| : || |||||  |||  |:: |||
Db      1 TGGATAAAAAATCWATWACACATGAAGAATTTATWMGACA 40


RESULT 100
US-09-605-192-4/c
; Sequence 4, Application US/09605192
; Patent No. 6323009
; GENERAL INFORMATION:
;  APPLICANT:  Dean, Frank B.
;  APPLICANT:  Laskin, Roger S.
;  APPLICANT:  Nelson, John
;  TITLE OF INVENTION:  Multiply-primed Amplification of Nucleic Acid Sequences
;  FILE REFERENCE:  469290-41
;  CURRENT APPLICATION NUMBER: US/09/605,192
;  CURRENT FILING DATE:  2000-06-28
;  NUMBER OF SEQ ID NOS:  9
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 40
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
;   OTHER INFORMATION: primer that anneals to M13 (+)-strand DNA.
US-09-605-192-4

  Query Match             0.8%;  Score 19.2;  DB 4;  Length 40;
  Best Local Similarity   67.5%;  Pred. No. 2.7e+04;
  Matches   27;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   2216 CCAGACAACTTCTGTTTCTTCTCTTGGTGAAATAATAAAA 2255
          ||   | | | ||  | ||||||||||  |||| || ||||
Db     40 CCCCGCTAATCCTAATCCTTCTCTTGAGGAAAAAAAAAAA 1


Search completed: April 19, 2003, 10:07:20
Job time : 102 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: April 19, 2003, 09:03:23 ; Search time 241 Seconds
(without alignments)
9778.848 Million cell updates/sec

Title: US-09-920-677-3
Perfect score: 2346
Sequence: 1 gcacgaggctgcggcgggtc.....gtaacccagctgcggagcct 2346

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 639749 seqs, 502280978 residues

Total number of hits satisfying chosen parameters: 386756

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	24.2	1.0	57	7	US-08-781-986A-1737	Sequence 1737, Ap
	2	23.6	1.0	60	10	US-09-922-261-231	Sequence 231, App
c	3	22.8	1.0	60	10	US-09-925-301-758	Sequence 758, App
	4	22.6	1.0	54	9	US-10-033-297-162	Sequence 162, App
	5	22.6	1.0	54	9	US-09-940-244-162	Sequence 162, App
	6	22.4	1.0	53	10	US-09-983-965-2161	Sequence 2161, Ap
c	7	22	0.9	39	9	US-09-846-430A-5	Sequence 5, Appli
	8	21.8	0.9	55	7	US-08-781-986A-1690	Sequence 1690, Ap
c	9	21.8	0.9	58	10	US-09-783-590-2406	Sequence 2406, Ap
	10	21.4	0.9	31	10	US-09-801-274-808	Sequence 808, App
	11	21.2	0.9	37	9	US-09-877-478-4141	Sequence 4141, Ap
	12	21.2	0.9	58	10	US-09-764-860-203	Sequence 203, App
	13	21	0.9	59	12	US-10-013-737-17	Sequence 17, Appl
	14	20.8	0.9	40	10	US-09-798-675-21	Sequence 21, Appl
c	15	20.8	0.9	40	10	US-09-798-675-22	Sequence 22, Appl
	16	20.6	0.9	34	9	US-09-377-885A-34	Sequence 34, Appl
c	17	20.6	0.9	49	9	US-09-931-325A-72	Sequence 72, Appl
	18	20.6	0.9	54	10	US-09-983-965-452	Sequence 452, App
c	19	20.6	0.9	55	9	US-09-931-325A-63	Sequence 63, Appl
c	20	20.6	0.9	55	9	US-09-931-325A-75	Sequence 75, Appl
	21	20.6	0.9	57	9	US-09-931-325A-71	Sequence 71, Appl
	22	20.2	0.9	54	7	US-08-781-986A-5031	Sequence 5031, Ap
	23	20	0.9	47	9	US-09-853-526-315	Sequence 315, App
	24	20	0.9	47	10	US-09-901-484A-315	Sequence 315, App
	25	20	0.9	49	9	US-10-085-906-261	Sequence 261, App
c	26	20	0.9	53	7	US-08-781-986A-2757	Sequence 2757, Ap
c	27	20	0.9	56	10	US-09-944-036-30	Sequence 30, Appl
	28	20	0.9	60	9	US-10-046-935-335	Sequence 335, App
c	29	20	0.9	60	9	US-10-057-940-14	Sequence 14, Appl
	30	20	0.9	60	9	US-09-878-178-335	Sequence 335, App
	31	20	0.9	60	9	US-10-079-178-8	Sequence 8, Appli
	32	20	0.9	60	9	US-10-146-502-335	Sequence 335, App
	33	20	0.9	60	10	US-09-983-965-83	Sequence 83, Appl
	34	19.8	0.8	48	10	US-09-922-261-233	Sequence 233, App
c	35	19.8	0.8	55	9	US-09-866-925-129	Sequence 129, App
c	36	19.8	0.8	59	9	US-10-085-906-282	Sequence 282, App
c	37	19.6	0.8	44	9	US-09-232-785-363	Sequence 363, App
c	38	19.6	0.8	46	10	US-09-827-289-22	Sequence 22, Appl
c	39	19.6	0.8	46	10	US-09-827-289-26	Sequence 26, Appl
c	40	19.6	0.8	53	9	US-09-938-842A-3026	Sequence 3026, Ap
c	41	19.6	0.8	57	9	US-10-046-935-2169	Sequence 2169, Ap
c	42	19.6	0.8	57	9	US-09-878-178-2169	Sequence 2169, Ap
c	43	19.6	0.8	57	9	US-10-127-427-7	Sequence 7, Appli
c	44	19.6	0.8	57	9	US-10-146-502-2169	Sequence 2169, Ap
c	45	19.6	0.8	57	10	US-09-333-527-7	Sequence 7, Appli
	46	19.4	0.8	54	9	US-10-085-906-66	Sequence 66, Appl
c	47	19.4	0.8	54	10	US-09-983-965-452	Sequence 452, App
	48	19.4	0.8	54	10	US-09-263-959-519	Sequence 519, App
c	49	19.4	0.8	60	9	US-10-046-935-335	Sequence 335, App
c	50	19.4	0.8	60	9	US-09-878-178-335	Sequence 335, App
	51	19.4	0.8	60	9	US-10-208-155-6	Sequence 6, Appli
c	52	19.4	0.8	60	9	US-10-146-502-335	Sequence 335, App
	53	19.2	0.8	37	9	US-09-730-289B-2736	Sequence 2736, Ap
c	54	19.2	0.8	39	10	US-09-263-959-213	Sequence 213, App
c	55	19.2	0.8	43	10	US-09-753-436-101	Sequence 101, App
c	56	19.2	0.8	46	10	US-09-263-959-148	Sequence 148, App
	57	19.2	0.8	48	9	US-09-999-724-14	Sequence 14, Appl
c	58	19.2	0.8	49	9	US-09-796-679-18	Sequence 18, Appl
c	59	19.2	0.8	49	9	US-10-085-906-204	Sequence 204, App
c	60	19.2	0.8	53	9	US-09-907-900-52	Sequence 52, Appl
c	61	19.2	0.8	53	9	US-09-907-719-52	Sequence 52, Appl
c	62	19.2	0.8	53	10	US-09-855-797A-52	Sequence 52, Appl
	63	19.2	0.8	54	10	US-09-970-308-8	Sequence 8, Appli
c	64	19.2	0.8	54	10	US-09-970-308-9	Sequence 9, Appli
	65	19.2	0.8	57	10	US-09-426-548-41	Sequence 41, Appl
	66	19.2	0.8	60	9	US-09-902-941-2002	Sequence 2002, Ap
	67	19.2	0.8	60	9	US-10-017-754-2002	Sequence 2002, Ap
c	68	19	0.8	46	9	US-10-085-906-255	Sequence 255, App
	69	19	0.8	46	10	US-09-827-289-20	Sequence 20, Appl
	70	19	0.8	46	10	US-09-827-289-24	Sequence 24, Appl
	71	19	0.8	53	9	US-09-938-842A-3904	Sequence 3904, Ap
c	72	19	0.8	54	10	US-09-263-959-519	Sequence 519, App
c	73	19	0.8	55	9	US-10-085-906-18	Sequence 18, Appl
c	74	19	0.8	57	9	US-10-007-132-59	Sequence 59, Appl
c	75	19	0.8	57	10	US-09-944-036-31	Sequence 31, Appl
c	76	19	0.8	57	10	US-09-943-286-7	Sequence 7, Appli
	77	19	0.8	60	9	US-09-954-692-17	Sequence 17, Appl
	78	19	0.8	60	10	US-09-559-671A-17	Sequence 17, Appl
	79	18.8	0.8	37	9	US-09-877-478-4121	Sequence 4121, Ap
	80	18.8	0.8	38	9	US-09-877-478-3202	Sequence 3202, Ap
	81	18.8	0.8	39	9	US-09-247-890-14	Sequence 14, Appl
c	82	18.8	0.8	42	9	US-10-043-573-131	Sequence 131, App
c	83	18.8	0.8	51	9	US-10-046-935-34	Sequence 34, Appl
c	84	18.8	0.8	51	9	US-09-878-178-34	Sequence 34, Appl
c	85	18.8	0.8	51	9	US-10-146-502-34	Sequence 34, Appl
	86	18.8	0.8	52	9	US-10-027-805-1	Sequence 1, Appli
c	87	18.8	0.8	52	9	US-09-907-900-58	Sequence 58, Appl
c	88	18.8	0.8	52	9	US-09-907-719-58	Sequence 58, Appl
	89	18.8	0.8	52	9	US-10-027-804-1	Sequence 1, Appli
c	90	18.8	0.8	52	10	US-09-855-797A-58	Sequence 58, Appl
	91	18.8	0.8	52	10	US-09-903-410-1	Sequence 1, Appli
	92	18.8	0.8	53	9	US-09-995-225-51	Sequence 51, Appl

93	18.8	0.8	53	9	US-10-083-168-92	Sequence 92, Appl
c 94	18.8	0.8	54	7	US-08-781-986A-2724	Sequence 2724, Ap
95	18.8	0.8	56	10	US-09-944-036-30	Sequence 30, Appl
96	18.8	0.8	57	7	US-08-781-986A-1941	Sequence 1941, Ap
c 97	18.8	0.8	57	9	US-09-981-803-29	Sequence 29, Appl
c 98	18.8	0.8	58	10	US-09-878-574-9023	Sequence 9023, Ap
c 99	18.8	0.8	59	9	US-10-007-968-8	Sequence 8, Appli
100	18.8	0.8	59	9	US-10-007-280A-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-08-781-986A-1737
; Sequence 1737, Application US/08781986A
; Publication No. US20030054436A1
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  1737:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  57 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-781-986A-1737

  Query Match             1.0%;  Score 24.2;  DB 7;  Length 57;
  Best Local Similarity   71.1%;  Pred. No. 7.4e+03;
  Matches   32;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    421 AGAAATGCTAAAGATACAGCTCATACAAAAGCAGAACGGAATATT 465
          | | | |||||||| | || | ||| || |||| ||  || ||||
Db      1 AAAGAAGCTAAAGAAAAAGTTGATAAAACAGCAAAAGTGATTATT 45


RESULT 2
US-09-922-261-231
; Sequence 231, Application US/09922261
; Patent No. US20020111471A1
; GENERAL INFORMATION:
;  APPLICANT: COGENT NEUROSCIENCE, Inc.
;  APPLICANT:  Lo, Donald C.
;  APPLICANT:  Barney, Shawn
;  APPLICANT:  Thomas, Mary Beth
;  APPLICANT:  Portbury, Stuart D.
;  APPLICANT:  Puranam, Kasturi
;  APPLICANT:  Katz, Lawrence C.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
;  TITLE OF INVENTION:  CELL DEATH
;  FILE REFERENCE: 10001-005-999
;  CURRENT APPLICATION NUMBER: US/09/922,261
;  CURRENT FILING DATE:  2001-08-03
;  PRIOR APPLICATION NUMBER: US/09/461,697
;  PRIOR FILING DATE: 1999-12-14
;  NUMBER OF SEQ ID NOS: 466
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 231
;    LENGTH: 60
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-922-261-231

  Query Match             1.0%;  Score 23.6;  DB 10;  Length 60;
  Best Local Similarity   64.8%;  Pred. No. 1.1e+04;
  Matches   35;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy   1635 AAAAAGGTGGAGAGGGAGATGTGTGAGCATCCTGCAAGGTGAAACAAGACTCAA 1688
          || ||| || ||  ||| ||| |  || | |  | || | |||| |||| | ||
Db      7 AAGAAGATGAAGGTGGAAATGAGGAAGAAGCTGGAAAAGAGAAAGAAGATTTAA 60


RESULT 3
US-09-925-301-758/c
; Sequence 758, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 758
;    LENGTH: 60
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (36)
;    OTHER INFORMATION: n equals a,t,g, or c
;    NAME/KEY: misc_feature
;    LOCATION: (38)
;    OTHER INFORMATION: n equals a,t,g, or c
;    NAME/KEY: misc_feature
;    LOCATION: (40)
;    OTHER INFORMATION: n equals a,t,g, or c
;    NAME/KEY: misc_feature
;    LOCATION: (45)
;    OTHER INFORMATION: n equals a,t,g, or c
;    NAME/KEY: misc_feature
;    LOCATION: (46)
;    OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-758

  Query Match             1.0%;  Score 22.8;  DB 10;  Length 60;
  Best Local Similarity   63.8%;  Pred. No. 1.8e+04;
  Matches   30;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   1756 TTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
          ||||||||||  ||  .  |      |||||||||   |||  |||| |
Db     56 TTTTAAAAAANNAAAANANANAAAAAAAAAAAAAAAAAAAAAAAAAA 10
```

```
RESULT 4
US-10-033-297-162
; Sequence 162, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;        APPLICANT: Hall, Jeff G.
;                   Lyamichev, Victor I.
;                   Mast, Andrea L.
;                   Brow, Mary Ann D.
;        TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                            Sequential Invasive Cleavages
;        NUMBER OF SEQUENCES: 163
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/033,297
;             FILING DATE: 12-No. US20020187486A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,597
;             FILING DATE: 09-Jul-1999
;             APPLICATION NUMBER: US/08/823,516
;             FILING DATE: 24-MAR-1997
;             APPLICATION NUMBER: PCT/US97/01072
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 162:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 54 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: other nucleic acid
;             DESCRIPTION: /desc = "DNA"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 162:
US-10-033-297-162

  Query Match             1.0%;  Score 22.6;  DB 9;  Length 54;
  Best Local Similarity   75.7%;  Pred. No. 1.9e+04;
  Matches   28;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy     128 AGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGA 164
           ||  |||||| | |||  || |||| |||| ||||||
Db       4 AGGAAGAGGAGGAGGGTGCTCAGGAGGAGCGGGAGGA 40


RESULT 5
US-09-940-244-162
; Sequence 162, Application US/09940244
; Publication No. US20030044796A1
; GENERAL INFORMATION:
;  APPLICANT: Neri, Bruce P.
;  APPLICANT:  Hall, Jeff G.
;  APPLICANT:  Lyamichev, Victor
;  APPLICANT:  Smith, Lloyd M.
;  TITLE OF INVENTION: Reactions on Dendrimers
;  FILE REFERENCE: FORS-06478
;  CURRENT APPLICATION NUMBER: US/09/940,244
;  CURRENT FILING DATE:  2002-05-06
;  NUMBER OF SEQ ID NOS: 422
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 162
;   LENGTH: 54
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
US-09-940-244-162

  Query Match             1.0%;  Score 22.6;  DB 9;  Length 54;
  Best Local Similarity   75.7%;  Pred. No. 1.9e+04;
  Matches   28;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy     128 AGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGA 164
           ||  |||||| | |||  || |||| |||| ||||||
Db       4 AGGAAGAGGAGGAGGGTGCTCAGGAGGAGCGGGAGGA 40


RESULT 6
US-09-983-965-2161
; Sequence 2161, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 37-21(10297)C
;  CURRENT APPLICATION NUMBER: US/09/983,965
;  CURRENT FILING DATE:  2001-10-26
;  PRIOR APPLICATION NUMBER: US 09/465,231
;  PRIOR FILING DATE: 1999-12-15
;  PRIOR APPLICATION NUMBER: US 60/113,678
;  PRIOR FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 2161
;   LENGTH: 53
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 40-LIB3057-003-Q1-K1-B8
US-09-983-965-2161

  Query Match             1.0%;  Score 22.4;  DB 10;  Length 53;
  Best Local Similarity   72.5%;  Pred. No. 2.1e+04;
  Matches   29;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy    1129 CCCTTTAAACCTCTGTTGCAATCTGAAGAGGATGTAAGTC 1168
           |||| || || ||| || ||| ||| |||| |  |||||
Db      11 CCCTGTACACATCTTTTTCAAACTGGAGAGCCTTCAAGTC 50


RESULT 7
US-09-846-430A-5/c
; Sequence 5, Application US/09846430A
; Publication No. US20030049620A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: LAI, Jennifer H.
;  APPLICANT:  PHILLIPS, Vincent
;  APPLICANT:  WATSON, Andrew R
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR POLYNUCLEOTIDE ANALYSIS
;  TITLE OF INVENTION:  USING GENERIC CAPTURE SEQUENCES
;  FILE REFERENCE: 5100-7001 / 0016-US
;  CURRENT APPLICATION NUMBER: US/09/846,430A
;  CURRENT FILING DATE:  2001-04-30
;  PRIOR APPLICATION NUMBER: 60/200,635
;  PRIOR FILING DATE: 2000-04-28
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 39
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  GC1a-B
US-09-846-430A-5

  Query Match             0.9%;  Score 22;  DB 9;  Length 39;
  Best Local Similarity   73.7%;  Pred. No. 2.2e+04;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1763 AAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAA 1800
          |||||| ||| |||  || |||||||  |||  ||||
Db     39 AAATCAGTCAGTGGCCCAGAAAAAAAAAAAAAAAAAA 2


RESULT 8
US-08-781-986A-1690
; Sequence 1690, Application US/08781986A
; Publication No. US20030054436A1
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  1690:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  55 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-781-986A-1690

  Query Match             0.9%;  Score 21.8;  DB 7;  Length 55;
  Best Local Similarity   70.7%;  Pred. No. 3.1e+04;
  Matches   29;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   2174 CCAAACTTTAAAATGCGAAATTATTGGTTGGTGTGAAGAAA 2214
          |||||||||| |||  |||||  |||||   |   |||| |
Db      1 CCAAACTTTATAATATGAAATGCTTGGTAATTACAAAGATA 41


RESULT 9
US-09-783-590-2406/c
; Sequence 2406, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
;  APPLICANT: Dillon, Patrick J.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Li, Haodong
;  APPLICANT:  Rosen, Craig A.
;  APPLICANT:  Ruben, Steven M.
;  TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
;  FILE REFERENCE: PO-16.2C1
;  CURRENT APPLICATION NUMBER: US/09/783,590
;  CURRENT FILING DATE:  2000-02-15
;  PRIOR APPLICATION NUMBER: 08/420,856
;  PRIOR FILING DATE: 1995-04-12
;  PRIOR APPLICATION NUMBER: 08/346,731
;  PRIOR FILING DATE: 1994-11-21
;  NUMBER OF SEQ ID NOS: 12485
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2406
;   LENGTH: 58
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (34)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (41)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (47)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (52)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (58)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-2406

  Query Match             0.9%;  Score 21.8;  DB 10;  Length 58;
  Best Local Similarity   64.4%;  Pred. No. 3.3e+04;
  Matches   29;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   1758 TTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
          || |||| |   | || ||   |||||||||   |||  |||| |
Db     54 TTNAAAANTTTTTNAAAGGGNAAAAAAAAAAAAAAAAAAAAAAAA 10


RESULT 10
US-09-801-274-808
; Sequence 808, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
```

```
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 808
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-808

  Query Match             0.9%;  Score 21.4;  DB 10;  Length 31;
  Best Local Similarity   88.0%;  Pred. No. 2.7e+04;
  Matches   22;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy   1777 TGCAAAAAAAAACTTAAAGCAAAAT 1801
          | ||||||||:||||| ||||||||
Db      6 TCCAAAAAAAARACTTACAGCAAAAT 30


RESULT 11
US-09-877-478-4141
; Sequence 4141, Application US/09877478
; Publication No. US20030068301A1
; GENERAL INFORMATION:
;  APPLICANT: Ribozyme Pharmaceuticals, Inc.
;  APPLICANT:  Draper, Kenneth
;  APPLICANT:  Blatt, Larry
;  APPLICANT:  McSwiggen, Jim
;  APPLICANT:  Morrissey, Dave
;  TITLE OF INVENTION: Method and Reagent for Inhibiting Hepatitis B Virus Replication
;  FILE REFERENCE: MBHB00-845-H (400/029)
;  CURRENT APPLICATION NUMBER: US/09/877,478
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: US 07/882,712
;  PRIOR FILING DATE: 1992-05-14
;  PRIOR APPLICATION NUMBER: US 09/531,025
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: US 09/636,385
;  PRIOR FILING DATE: 2000-08-09
;  PRIOR APPLICATION NUMBER: US 09/696,347
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 08/193,627
;  PRIOR FILING DATE: 1994-02-07
;  PRIOR APPLICATION NUMBER: US 08/433,993
;  PRIOR FILING DATE: 1995-05-04
;  PRIOR APPLICATION NUMBER: US 08/434,504
;  PRIOR FILING DATE: 1995-05-04
;  PRIOR APPLICATION NUMBER: US 09/436,430
;  PRIOR FILING DATE: 1999-11-08
;  NUMBER OF SEQ ID NOS: 6586
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4141
;   LENGTH: 37
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Enzymatic Nucleic Acid
US-09-877-478-4141

  Query Match             0.9%;  Score 21.2;  DB 9;  Length 37;
  Best Local Similarity   58.8%;  Pred. No. 3.5e+04;
  Matches   20;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    175 TTAAATGAAAGCATGGACCATGGGGGAGTTGGAC 208
          :: ||:||  |||:| || |:| | |||: ||||
Db      4 UUGAAUGAUGGCAUGCACUAUGCGCGAGUAGGAC 37


RESULT 12
US-09-764-860-203
; Sequence 203, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC008
;  CURRENT APPLICATION NUMBER: US/09/764,860
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1198
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 203
;   LENGTH: 58
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-860-203

  Query Match             0.9%;  Score 21.2;  DB 10;  Length 58;
  Best Local Similarity   64.0%;  Pred. No. 4.7e+04;
  Matches   32;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy   2089 TACGTGCAAACAACCTGAATCTTTTTTTTATATAAATATATATTTTTCAA 2138
          ||  | |||||||   ||| |||||||||| | |   | | ||   || ||
Db      8 TATTTTCAAACAATAGGAAACTTTTTTTTTTTTTTTTTTTAAAATTTAA 57


RESULT 13
US-10-013-737-17
; Sequence 17, Application US/10013737
; Patent No. US20020115093A1
; GENERAL INFORMATION:
;  APPLICANT: Reynolds, Mark A.
;  APPLICANT:  Ruvolo, Michael
;  APPLICANT:  Arnold, Jr., Lyle J.
;  TITLE OF INVENTION: COMBINED POLYNUCLEOTIDE SEQUENCES AS DISCRETE ASSAY
;  TITLE OF INVENTION:  ENDPOINTS
;  FILE REFERENCE: IN-0017 US
;  CURRENT APPLICATION NUMBER: US/10/013,737
;  CURRENT FILING DATE:  2001-11-12
;  PRIOR APPLICATION NUMBER: US/09/532,656
;  PRIOR FILING DATE: 2000-03-20
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PERL Program
; SEQ ID NO 17
;   LENGTH: 59
;   TYPE: DNA
;   ORGANISM: Candida albicans
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020115093A1 F23S805
US-10-013-737-17

  Query Match             0.9%;  Score 21;  DB 12;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 5.4e+04;
  Matches   30;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   1078 TTCTTTAGACACATTAACTGGGAAGAACTTCTGGCTCGAAAGGTG 1122
          ||| ||||  ||||||| |   | ||||  ||  | |||| |||
Db      5 TTCATTAGTGACATTAAAAGAAGAAAACTGATGTTTTGAAATGTG 49


RESULT 14
US-09-798-675-21
; Sequence 21, Application US/09798675
; Patent No. US20020106798A1
; GENERAL INFORMATION:
;  APPLICANT: Emory University
;  TITLE OF INVENTION: HIV VACCINES
;  FILE REFERENCE: E056 2020
;  CURRENT APPLICATION NUMBER: US/09/798,675
;  CURRENT FILING DATE:  2001-12-11
;  PRIOR APPLICATION NUMBER: US 60/186,364
;  PRIOR FILING DATE: 2000-03-02
;  PRIOR APPLICATION NUMBER: US 60/251,083
;  PRIOR FILING DATE: 2000-12-01
;  NUMBER OF SEQ ID NOS: 24
```

```
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
;   LENGTH: 40
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: E478Q RT5 primer
US-09-798-675-21

  Query Match             0.9%;  Score 20.8;  DB 10;  Length 40;
  Best Local Similarity   70.0%;  Pred. No. 4.7e+04;
  Matches   28;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    967 CCCTACCTCACACAAGAAGCCAGAGATCTGCTTAAAAAGC 1006
          ||||| || |||||| ||  |||| | ||    | | ||||
Db      1 CCCTAACTAACACAACAAATCAGAAAACTCAGTTACAAGC 40


RESULT 15
US-09-798-675-22/c
; Sequence 22, Application US/09798675
; Patent No. US20020106798A1
; GENERAL INFORMATION:
;  APPLICANT: Emory University
;  TITLE OF INVENTION: HIV VACCINES
;  FILE REFERENCE: E056 2020
;  CURRENT APPLICATION NUMBER: US/09/798,675
;  CURRENT FILING DATE:  2001-12-11
;  PRIOR APPLICATION NUMBER: US 60/186,364
;  PRIOR FILING DATE: 2000-03-02
;  PRIOR APPLICATION NUMBER: US 60/251,083
;  PRIOR FILING DATE: 2000-12-01
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
;   LENGTH: 40
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: E478Q RT6 primer
US-09-798-675-22

  Query Match             0.9%;  Score 20.8;  DB 10;  Length 40;
  Best Local Similarity   70.0%;  Pred. No. 4.7e+04;
  Matches   28;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    967 CCCTACCTCACACAAGAAGCCAGAGATCTGCTTAAAAAGC 1006
          ||||| || |||||| ||  |||| | ||    | | ||||
Db     40 CCCTAACTAACACAACAAATCAGAAAACTCAGTTACAAGC 1


RESULT 16
US-09-377-885A-34
; Sequence 34, Application US/09377885A
; Publication No. US20030050258A1
; GENERAL INFORMATION:
;  APPLICANT: Calos, Michele P.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR GENOMIC MODIFICATION
;  FILE REFERENCE: STAN-301
;  CURRENT APPLICATION NUMBER: US/09/377,885A
;  CURRENT FILING DATE:  2003-02-07
;  PRIOR APPLICATION NUMBER: 60/097,166
;  PRIOR FILING DATE: 1998-08-19
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34
;   LENGTH: 34
;   TYPE: DNA
;   ORGANISM: Mus sp.
US-09-377-885A-34

  Query Match             0.9%;  Score 20.6;  DB 9;  Length 34;
  Best Local Similarity   85.2%;  Pred. No. 4.7e+04;
  Matches   23;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy   2107 ATCTTTTTTTTATATAAATATATATTT 2133
          || || | |||||||| ||||||||||
Db      6 ATATTATATTTATATATATATATATTT 32


RESULT 17
US-09-931-325A-72/c
; Sequence 72, Application US/09931325A
; Publication No. US20030054337A1
; GENERAL INFORMATION:
;  APPLICANT: Birkett, Ashley J.
;  TITLE OF INVENTION: MALARIA IMMUNOGEN AND VACCINE
;  FILE REFERENCE: 4564/83503 ICC-103.1
;  CURRENT APPLICATION NUMBER: US/09/931,325A
;  CURRENT FILING DATE:  2002-02-22
;  PRIOR APPLICATION NUMBER: 60/225,843
;  PRIOR FILING DATE: 2000-08-16
;  PRIOR APPLICATION NUMBER: USSN NOT YET ASSIGND
;  PRIOR FILING DATE: 2001-08-15
;  NUMBER OF SEQ ID NOS: 186
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 72
;   LENGTH: 49
;   TYPE: DNA
;   ORGANISM: Plasmodium falciparum
US-09-931-325A-72

  Query Match             0.9%;  Score 20.6;  DB 9;  Length 49;
  Best Local Similarity   67.4%;  Pred. No. 6e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    423 AAATGCTAAAGATACAGCTCATACAAAAGCAGAACGGAATATT 465
          |||||| ||   ||  ||| || |||| ||  | | |||| ||
Db     44 AAATGCCAACCCTAACGCTAATCCAAACGCCAACCCGAATGTT 2


RESULT 18
US-09-983-965-452
; Sequence 452, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 37-21(10297)C
;  CURRENT APPLICATION NUMBER: US/09/983,965
;  CURRENT FILING DATE:  2001-10-26
;  PRIOR APPLICATION NUMBER: US 09/465,231
;  PRIOR FILING DATE: 1999-12-15
;  PRIOR APPLICATION NUMBER: US 60/113,678
;  PRIOR FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 452
;   LENGTH: 54
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 53-BOVMS1-015-Q1-E1-F10
US-09-983-965-452

  Query Match             0.9%;  Score 20.6;  DB 10;  Length 54;
  Best Local Similarity   67.4%;  Pred. No. 6.5e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1760 AAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
          ||||||  ||  ||    | |||||||||| | |||  |||| |
Db      1 AAAAAAAAAAAAAAAAAAGAAAAAAAAAAATAAAAAAAAAAAA 43
```

```
RESULT 19
US-09-931-325A-63/c
; Sequence 63, Application US/09931325A
; Publication No. US20030054337A1
; GENERAL INFORMATION:
;  APPLICANT: Birkett, Ashley J.
;  TITLE OF INVENTION: MALARIA IMMUNOGEN AND VACCINE
;  FILE REFERENCE: 4564/83503 ICC-103.1
;  CURRENT APPLICATION NUMBER: US/09/931,325A
;  CURRENT FILING DATE:  2002-02-22
;  PRIOR APPLICATION NUMBER: 60/225,843
;  PRIOR FILING DATE: 2000-08-16
;  PRIOR APPLICATION NUMBER: USSN NOT YET ASSIGND
;  PRIOR FILING DATE: 2001-08-15
;  NUMBER OF SEQ ID NOS: 186
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
;   LENGTH: 55
;   TYPE: DNA
;   ORGANISM: Plasmodium falciparum
US-09-931-325A-63

  Query Match             0.9%;  Score 20.6;  DB 9;  Length 55;
  Best Local Similarity   67.4%;  Pred. No. 6.5e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy     423 AAATGCTAAAGATACAGCTCATACAAAAGCAGAACGGAATATT 465
           |||||| ||    ||  ||| || |||| ||  | | |||| ||
Db      44 AAATGCCAACCCTAACGCTAATCCAAACGCCAACCCGAATGTT 2


RESULT 20
US-09-931-325A-75/c
; Sequence 75, Application US/09931325A
; Publication No. US20030054337A1
; GENERAL INFORMATION:
;  APPLICANT: Birkett, Ashley J.
;  TITLE OF INVENTION: MALARIA IMMUNOGEN AND VACCINE
;  FILE REFERENCE: 4564/83503 ICC-103.1
;  CURRENT APPLICATION NUMBER: US/09/931,325A
;  CURRENT FILING DATE:  2002-02-22
;  PRIOR APPLICATION NUMBER: 60/225,843
;  PRIOR FILING DATE: 2000-08-16
;  PRIOR APPLICATION NUMBER: USSN NOT YET ASSIGND
;  PRIOR FILING DATE: 2001-08-15
;  NUMBER OF SEQ ID NOS: 186
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
;   LENGTH: 55
;   TYPE: DNA
;   ORGANISM: Plasmodium falciparum
US-09-931-325A-75

  Query Match             0.9%;  Score 20.6;  DB 9;  Length 55;
  Best Local Similarity   67.4%;  Pred. No. 6.5e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy     423 AAATGCTAAAGATACAGCTCATACAAAAGCAGAACGGAATATT 465
           |||||| ||    ||  ||| || |||| ||  | | |||| ||
Db      50 AAATGCCAACCCTAACGCTAATCCAAACGCCAACCCGAATGTT 8


RESULT 21
US-09-931-325A-71
; Sequence 71, Application US/09931325A
; Publication No. US20030054337A1
; GENERAL INFORMATION:
;  APPLICANT: Birkett, Ashley J.
;  TITLE OF INVENTION: MALARIA IMMUNOGEN AND VACCINE
;  FILE REFERENCE: 4564/83503 ICC-103.1
;  CURRENT APPLICATION NUMBER: US/09/931,325A
;  CURRENT FILING DATE:  2002-02-22
;  PRIOR APPLICATION NUMBER: 60/225,843
;  PRIOR FILING DATE: 2000-08-16
;  PRIOR APPLICATION NUMBER: USSN NOT YET ASSIGND
;  PRIOR FILING DATE: 2001-08-15
;  NUMBER OF SEQ ID NOS: 186
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 71
;   LENGTH: 57
;   TYPE: DNA
;   ORGANISM: Plasmodium falciparum
US-09-931-325A-71

  Query Match             0.9%;  Score 20.6;  DB 9;  Length 57;
  Best Local Similarity   67.4%;  Pred. No. 6.7e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy     423 AAATGCTAAAGATACAGCTCATACAAAAGCAGAACGGAATATT 465
           |||||| ||    ||  ||| || |||| ||  | | |||| ||
Db      10 AAATGCCAACCCTAACGCTAATCCAAACGCCAACCCGAATGTT 52


RESULT 22
US-08-781-986A-5031
; Sequence 5031, Application US/08781986A
; Publication No. US20030054436A1
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  5031:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  54 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-781-986A-5031

  Query Match             0.9%;  Score 20.2;  DB 7;  Length 54;
  Best Local Similarity   75.8%;  Pred. No. 8.3e+04;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    2106 AATCTTTTTTTTATATAAATATATATTTTTCAA 2138
           ||||| | |||||  ||  ||| || |||||||
Db      11 AATCTATATTTTACTTACTTATCTAGTTTTCAA 43


RESULT 23
```

```
US-09-853-526-315
; Sequence 315, Application US/09853526
; Patent No. US20020165345A1
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Ilya, Chumakov
;  APPLICANT:  Bougueleret, Lydie
;  TITLE OF INVENTION: PROSTATE CANCER GENE
;  FILE REFERENCE: GENSET.18CP1CP
;  CURRENT APPLICATION NUMBER: US/09/853,526
;  CURRENT FILING DATE:  2001-05-11
;  PRIOR APPLICATION NUMBER: 09/338,907
;  PRIOR FILING DATE: 1999-06-23
;  PRIOR APPLICATION NUMBER: 08/996,306
;  PRIOR FILING DATE: 1997-12-22
;  PRIOR APPLICATION NUMBER: 60/099,658
;  PRIOR FILING DATE: 1998-09-09
;  PRIOR APPLICATION NUMBER: 09/218,207
;  PRIOR FILING DATE: 1998-12-22
;  NUMBER OF SEQ ID NOS: 578
;  SOFTWARE: Patent.pm
; SEQ ID NO 315
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 1..47
;   OTHER INFORMATION: polymorphic fragment 99-147-181, variant version of SEQ ID238
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: base G ; A in SEQ ID238
;   NAME/KEY: primer_bind
;   LOCATION: 1..23
;   OTHER INFORMATION: potential microsequencing oligo 99-147-181.mis1
;   NAME/KEY: primer_bind
;   LOCATION: 25..47
;   OTHER INFORMATION: complement potential microsequencing oligo 99-147-181.mis2
US-09-853-526-315

  Query Match             0.9%;  Score 20;  DB 9;  Length 47;
  Best Local Similarity   72.2%;  Pred. No. 8.5e+04;
  Matches   26;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy    397 GTGCTTAAAAAGGCAATGATAGTAAGAAATGCTAAA 432
          ||  | ||||||   ||||||| ||||||  || ||
Db      3 GTCATGAAAAAGAGCATGATAGAAAGAAAAACTTAA 38


RESULT 24
US-09-901-484A-315
; Sequence 315, Application US/09901484A
; Patent No. US20020119460A1
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Bougueleret, Lydie
;  TITLE OF INVENTION: Prostate Cancer Gene
;  FILE REFERENCE: GEN-T111XC3D2
;  CURRENT APPLICATION NUMBER: US/09/901,484A
;  CURRENT FILING DATE:  2001-07-09
;  PRIOR APPLICATION NUMBER: US 08/996,306
;  PRIOR FILING DATE: 1997-12-22
;  PRIOR APPLICATION NUMBER: US 60/099,658
;  PRIOR FILING DATE: 1998-09-09
;  PRIOR APPLICATION NUMBER: US 09/218,207
;  PRIOR FILING DATE: 1998-12-22
;  PRIOR APPLICATION NUMBER: US 09/338,907
;  PRIOR FILING DATE: 1999-06-23
;  PRIOR APPLICATION NUMBER: US 09/853,526
;  PRIOR FILING DATE: 2001-05-11
;  NUMBER OF SEQ ID NOS: 578
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 315
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1)..(47)
;   OTHER INFORMATION: polymorphic fragment 99-147-181, variant version of SEQ ID 238
;   NAME/KEY: allele
;   LOCATION: (24)..(24)
;   OTHER INFORMATION: polymorphic base G; A in SEQ ID 238
;   NAME/KEY: primer_bind
;   LOCATION: (1)..(23)
;   OTHER INFORMATION: potential microsequencing oligo 99-147-181.mis1
;   NAME/KEY: primer_bind
;   LOCATION: (25)..(47)
;   OTHER INFORMATION: complement potential microsequencing oligo 99-147-181.mis2
US-09-901-484A-315

  Query Match             0.9%;  Score 20;  DB 10;  Length 47;
  Best Local Similarity   72.2%;  Pred. No. 8.5e+04;
  Matches   26;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy    397 GTGCTTAAAAAGGCAATGATAGTAAGAAATGCTAAA 432
          ||  | ||||||  ||||||| ||||||  || ||
Db      3 GTCATGAAAAAGAGCATGATAGAAAGAAAAACTTAA 38


RESULT 25
US-10-085-906-261
; Sequence 261, Application US/10085906
; Publication No. US20030054371A1
; GENERAL INFORMATION:
;  APPLICANT: Ying, Vincent
;  APPLICANT:  Wu, Paul
;  APPLICANT:  Gray, Gary S.
;  TITLE OF INVENTION: POLYMORPHIC ELEMENTS IN THE
;  TITLE OF INVENTION:  COSTIMULATORY RECEPTOR LOCUS AND USES THEREOF
;  FILE REFERENCE: GNN-5343CP2
;  CURRENT APPLICATION NUMBER: US/10/085,906
;  CURRENT FILING DATE:  2002-02-27
;  PRIOR APPLICATION NUMBER: US 60/126,215
;  PRIOR FILING DATE: 1999-03-25
;  PRIOR APPLICATION NUMBER: US 09/534,061
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: PCT/US00/07938
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 545
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 261
;   LENGTH: 49
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-085-906-261

  Query Match             0.9%;  Score 20;  DB 9;  Length 49;
  Best Local Similarity   65.9%;  Pred. No. 8.7e+04;
  Matches   29;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2111 TTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCA 2154
          ||| ||||| ||  ||| ||||| |  | |   || || ||| |
Db      3 TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 46


RESULT 26
US-08-781-986A-2757/c
; Sequence 2757, Application US/08781986A
; Publication No. US20030054436A1
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
```

```
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  2757:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  53 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-781-986A-2757

  Query Match             0.9%;  Score 20;  DB 7;  Length 53;
  Best Local Similarity   61.5%;  Pred. No. 9.2e+04;
  Matches   32;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy     375 TGGGAAAATATTTGCCATGAAGGTGCTTAAAAAGGCAATGATAGTAAGAAAT 426
           || | || |  |  | || | || ||| ||| |||| | ||| ||| | |
Db      53 TGTGTAAGTTGTGTCTATAACGGAGCTAAAATAGGCGTTAATACGAAGCATT 2


RESULT 27
US-09-944-036-30/c
; Sequence 30, Application US/09944036
; Patent No. US20020055095A1
; GENERAL INFORMATION:
;  APPLICANT: YANG, Yeasing Y.
;  APPLICANT:  BRENTANO, Steven T.
;  APPLICANT:  BABOLA, Odile
;  APPLICANT:  TRAN, Nathalie
;  APPLICANT:  VERNET, Guy
;  TITLE OF INVENTION: AMPLIFICATION OF HIV-1 SEQUENCES FOR DETECTION OF
;  TITLE OF INVENTION:  SEQUENCES ASSOCIATED WITH DRUG-RESISTANCE MUTATIONS
;  FILE REFERENCE: GP114-02.UT
;  CURRENT APPLICATION NUMBER: US/09/944,036
;  CURRENT FILING DATE:  2001-08-31
;  PRIOR APPLICATION NUMBER: US 60/229,790
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
;   LENGTH: 56
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Capture
;   OTHER INFORMATION: oligonucleotide
US-09-944-036-30

  Query Match             0.9%;  Score 20;  DB 10;  Length 56;
  Best Local Similarity   72.2%;  Pred. No. 9.6e+04;
  Matches   26;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy    2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGA 2143
           | |||||||| | |   | | | ||||| |||||||
Db      54 TTTTTTTTTTTTTTTTTTTTTTTTTTAAAATAGA 19


RESULT 28
US-10-046-935-335
; Sequence 335, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527C1
;  CURRENT APPLICATION NUMBER: US/10/046,935
;  CURRENT FILING DATE:  2002-01-15
;  NUMBER OF SEQ ID NOS: 2239
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 335
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 10, 13
;   OTHER INFORMATION: n = A,T,C or G
US-10-046-935-335

  Query Match             0.9%;  Score 20;  DB 9;  Length 60;
  Best Local Similarity   59.3%;  Pred. No. 1e+05;
  Matches   32;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy    2098 ACAACCTGAATCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGAT 2151
           |||| ||     |||||||| | |   | | |||||||    |   |||| | |
Db       3 ACAAGCTNTTNTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTGGTT 56


RESULT 29
US-10-057-940-14/c
; Sequence 14, Application US/10057940
; Patent No. US20020168686A1
; GENERAL INFORMATION:
;  APPLICANT: Pantoliano, Michael W.
;  APPLICANT:  Salemme, F. Raymond
;  APPLICANT:  Carver, Jr., Theodore, E.
;  TITLE OF INVENTION: High Throughput Method for Functionally Classifying Proteins
;  TITLE OF INVENTION:  Identified Using a Genomics Approach
;  FILE REFERENCE: 1503.0310002/JAG/JSO
;  CURRENT APPLICATION NUMBER: US/10/057,940
;  CURRENT FILING DATE:  2002-05-03
;  PRIOR APPLICATION NUMBER: 09/190,128
;  PRIOR FILING DATE: 1998-11-12
;  PRIOR APPLICATION NUMBER: US 60/065,129
;  PRIOR FILING DATE: 1997-11-12
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: T- and A-rich tract with single-stranded and duplex DNA
US-10-057-940-14

  Query Match             0.9%;  Score 20;  DB 9;  Length 60;
  Best Local Similarity   65.9%;  Pred. No. 1e+05;
```

Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Qy   1756 TTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAA 1799
          ||||||||||  ||  ||       |||||||||   |||  |||
Db     44 TTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1
```

RESULT 30
US-09-878-178-335
; Sequence 335, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527
; CURRENT APPLICATION NUMBER: US/09/878,178
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 2237
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 335
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(60)
; OTHER INFORMATION: n = A,T,C or G
US-09-878-178-335

Query Match 0.9%; Score 20; DB 9; Length 60;
Best Local Similarity 59.3%; Pred. No. 1e+05;
Matches 32; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy   2098 ACAACCTGAATCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGAT 2151
          ||||  ||     ||||||||  |  |    |  |  |||||||     |   ||||  |  |
Db      3 ACAAGCTNTTNTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTGGTT 56
```

RESULT 31
US-10-079-178-8
; Sequence 8, Application US/10079178
; Publication No. US20030044816A1
; GENERAL INFORMATION:
; APPLICANT: Denison, Timothy J.
; APPLICANT: Sauer, Alexis
; APPLICANT: Golovchenko, Jene
; APPLICANT: Meller, Amit
; APPLICANT: Brandin, Eric
; APPLICANT: Branton, Daniel
; TITLE OF INVENTION: Characterization of Individual Polymer Molecules Based on
; TITLE OF INVENTION: Monomer-Interface Interactions
; FILE REFERENCE: 42697.142US4
; CURRENT APPLICATION NUMBER: US/10/079,178
; CURRENT FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 09/457,959
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: US 09/098,142
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: US 08/405,735
; PRIOR FILING DATE: 1995-03-17
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-10-079-178-8

Query Match 0.9%; Score 20; DB 9; Length 60;
Best Local Similarity 58.3%; Pred. No. 1e+05;
Matches 35; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

```
Qy   1741 AAAAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAA 1800
          |||||  |||     |    ||||||  ||  ||      |||||||||   |||  ||||
Db      1 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 60
```

RESULT 32
US-10-146-502-335
; Sequence 335, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C2
; CURRENT APPLICATION NUMBER: US/10/146,502
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 335
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10, 13
; OTHER INFORMATION: n = A,T,C or G
US-10-146-502-335

Query Match 0.9%; Score 20; DB 9; Length 60;
Best Local Similarity 59.3%; Pred. No. 1e+05;
Matches 32; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy   2098 ACAACCTGAATCTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGAT 2151
          ||||  ||     ||||||||  |  |    |  |  |||||||     |   ||||  |  |
Db      3 ACAAGCTNTTNTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTGGTT 56
```

RESULT 33
US-09-983-965-83
; Sequence 83, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 83
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 10-BOVMS1-018-Q1-E1-C5
US-09-983-965-83

Query Match 0.9%; Score 20; DB 10; Length 60;

Best Local Similarity 58.3%; Pred. No. 1e+05;
Matches 35; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

```
Qy   1741 AAAAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAA 1800
          ||||| |||     |    ||||||  ||  ||      |||||||||   |||  ||||
Db      1 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 60
```

RESULT 34
US-09-922-261-233
; Sequence 233, Application US/09922261
; Patent No. US20020111471A1
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/922,261
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US/09/461,697
; PRIOR FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 233
; LENGTH: 48
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-261-233

Query Match 0.8%; Score 19.8; DB 10; Length 48;
Best Local Similarity 63.8%; Pred. No. 9.7e+04;
Matches 30; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```
Qy   1642 TGGAGAGGGAGATGTGTGAGCATCCTGCAAGGTGAAACAAGACTCAA 1688
          || ||  ||| ||| |  || | |  | || | |||| |||| | ||
Db      2 TGAAGGTGGAAATGAGGAAGAAGCTGGAAAAGAGAAAGAAGATTTAA 48
```

RESULT 35
US-09-866-925-129/c
; Sequence 129, Application US/09866925
; Publication No. US20030039965A1
; GENERAL INFORMATION:
; APPLICANT: Feldmann, Richard J.
; TITLE OF INVENTION: ALGORITHMIC DETERMINATION OF FLANKING DNA SEQUENCES THAT
; TITLE OF INVENTION: CONTROL THE EXPRESSION OF SETS OF GENES IN PROKARYOTIC,
; TITLE OF INVENTION: ARCHEA AND EUKARYOTIC GENOMES
; FILE REFERENCE: 3124-Z
; CURRENT APPLICATION NUMBER: US/09/866,925
; CURRENT FILING DATE: 2001-05-30
; NUMBER OF SEQ ID NOS: 249
; SOFTWARE: Proprietary
; SEQ ID NO 129
; LENGTH: 55
; TYPE: DNA
; ORGANISM: A. Thaliania
; FEATURE:
; LOCATION: (499376)...(499430)
; OTHER INFORMATION: Chromosome = 2 Strand = positive ConnectronObjectNumber = 430
US-09-866-925-129

Query Match 0.8%; Score 19.8; DB 9; Length 55;
Best Local Similarity 63.8%; Pred. No. 1.1e+05;
Matches 30; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```
Qy   2089 TACGTGCAAACAACCTGAATCTTTTTTTTATATAAATATATATTTTT 2135
          || | | ||| |  || | | |||||||   || |  ||||  ||||
Db     54 TATGCGAAAAAACTCTAATTATTTTTTTATAATGATGATATTATTTT 8
```

RESULT 36
US-10-085-906-282/c
; Sequence 282, Application US/10085906
; Publication No. US20030054371A1
; GENERAL INFORMATION:
; APPLICANT: Ying, Vincent
; APPLICANT: Wu, Paul
; APPLICANT: Gray, Gary S.
; TITLE OF INVENTION: POLYMORPHIC ELEMENTS IN THE
; TITLE OF INVENTION: COSTIMULATORY RECEPTOR LOCUS AND USES THEREOF
; FILE REFERENCE: GNN-5343CP2
; CURRENT APPLICATION NUMBER: US/10/085,906
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 60/126,215
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: US 09/534,061
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: PCT/US00/07938
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 545
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 282
; LENGTH: 59
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-085-906-282

Query Match 0.8%; Score 19.8; DB 9; Length 59;
Best Local Similarity 60.0%; Pred. No. 1.1e+05;
Matches 33; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy    420 AAGAAATGCTAAAGATACAGCTCATACAAAAGCAGAACGGAATATTCTGGAGGAA 474
          ||||||     |||| ||||   |   ||||   |||   |||    ||||||||
Db     58 AAGAAAAAGAGAAGAGACAGAAGAGGGAAAAAGTGAAAAAAATGAGATGGAGGAA 4
```

RESULT 37
US-09-232-785-363/c
; Sequence 363, Application US/09232785
; Publication No. US20030049612A1
; GENERAL INFORMATION:
; APPLICANT: International Paper Co.
; APPLICANT: Echt, Craig. S
; APPLICANT: Nelson, C. Dana
; TITLE OF INVENTION: MICROSATELITE DNA MARKERS AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 4481/1E188US1
; CURRENT APPLICATION NUMBER: US/09/232,785
; CURRENT FILING DATE: 1999-01-19
; PRIOR APPLICATION NUMBER: 09/232,884
; PRIOR FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 397
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 363
; LENGTH: 44
; TYPE: DNA
; ORGANISM: Pinus taeda L.
US-09-232-785-363

Query Match 0.8%; Score 19.6; DB 9; Length 44;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 28; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```
Qy   2111 TTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
          ||| ||||| ||  ||| ||||| |  | |    || || |||
Db     43 TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATT 2
```

RESULT 38

```
US-09-827-289-22/c
; Sequence 22, Application US/09827289
; Patent No. US20020009716A1
; GENERAL INFORMATION:
;  APPLICANT: Abarzua, Patricio
;  TITLE OF INVENTION: Process for Allele Discrimination Using Primer
;  TITLE OF INVENTION:  Extension
;  FILE REFERENCE: 469290-55
;  CURRENT APPLICATION NUMBER: US/09/827,289
;  CURRENT FILING DATE:  2001-04-05
;  PRIOR APPLICATION NUMBER: U.S. 60/194843
;  PRIOR FILING DATE: 2000-04-05
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
;   LENGTH: 46
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: P1 primer for
;   OTHER INFORMATION: use in allele discrimination
US-09-827-289-22

  Query Match             0.8%;  Score 19.6;  DB 10;  Length 46;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+05;
  Matches   28;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    913 GAGAATAGAAAGAAAACAATTGACAAAATCCTCAAATGTAAA 954
          |||||||||| |||| ||  | |||| |||  |||
Db     45 GAGAATAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4


RESULT 39
US-09-827-289-26/c
; Sequence 26, Application US/09827289
; Patent No. US20020009716A1
; GENERAL INFORMATION:
;  APPLICANT: Abarzua, Patricio
;  TITLE OF INVENTION: Process for Allele Discrimination Using Primer
;  TITLE OF INVENTION:  Extension
;  FILE REFERENCE: 469290-55
;  CURRENT APPLICATION NUMBER: US/09/827,289
;  CURRENT FILING DATE:  2001-04-05
;  PRIOR APPLICATION NUMBER: U.S. 60/194843
;  PRIOR FILING DATE: 2000-04-05
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
;   LENGTH: 46
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: P1 primer for
;   OTHER INFORMATION: use in allele discrimination
US-09-827-289-26

  Query Match             0.8%;  Score 19.6;  DB 10;  Length 46;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+05;
  Matches   28;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    913 GAGAATAGAAAGAAAACAATTGACAAAATCCTCAAATGTAAA 954
          |||||||||| |||| ||  | |||| |||  |||
Db     45 GAGAATAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4


RESULT 40
US-09-938-842A-3026/c
; Sequence 3026, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 3026
;   LENGTH: 53
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-3026

  Query Match             0.8%;  Score 19.6;  DB 9;  Length 53;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+05;
  Matches   28;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    562 GAACTATTTATGCAGTTAGAAAGAGAGGGAATATTTATGGAA 603
          |||  ||||||| | |  | |||| | | || ||  || |||
Db     44 GAATCATTTATGTATTGGGTAAGAAAAGAAAAATAAATAGAA 3


RESULT 41
US-10-046-935-2169/c
; Sequence 2169, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527C1
;  CURRENT APPLICATION NUMBER: US/10/046,935
;  CURRENT FILING DATE:  2002-01-15
;  NUMBER OF SEQ ID NOS: 2239
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2169
;   LENGTH: 57
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 13
;   OTHER INFORMATION: n = A,T,C or G
US-10-046-935-2169

  Query Match             0.8%;  Score 19.6;  DB 9;  Length 57;
  Best Local Similarity   65.1%;  Pred. No. 1.2e+05;
  Matches   28;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   1760 AAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
          ||||||  ||  ||  | | |||||||||   |||  |||  |
Db     54 AAAAAAAAAAAAAAAAGGGAAAAAAAAAAAAAAAAAAAAGNA 12


RESULT 42
US-09-878-178-2169/c
; Sequence 2169, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
```

```
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527
;  CURRENT APPLICATION NUMBER: US/09/878,178
;  CURRENT FILING DATE:  2001-06-08
;  NUMBER OF SEQ ID NOS: 2237
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2169
;   LENGTH: 57
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(57)
;   OTHER INFORMATION: n = A,T,C or G
US-09-878-178-2169

  Query Match             0.8%;  Score 19.6;  DB 9;  Length 57;
  Best Local Similarity   65.1%;  Pred. No. 1.2e+05;
  Matches   28;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   1760 AAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
          ||||||  ||  ||  | | |||||||||   |||  |||  |
Db     54 AAAAAAAAAAAAAAAAGGGAAAAAAAAAAAAAAAAAAAAAGNA 12


RESULT 43
US-10-127-427-7/c
; Sequence 7, Application US/10127427
; Publication No. US20030051275A1
;   GENERAL INFORMATION:
;        APPLICANT: Paul CHRISTOU; Eva STROGER; Rainer FISCHER; Carmen MARTIN-VAQUERO;
;                   Stefan SC
;        TITLE OF INVENTION: METHODS AND MEANS FOR EXPRESSION OF MAMMALIAN
;                            POLYPEPTIDES
;        NUMBER OF SEQUENCES: 43
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fulbright & Jaworski L.L.P.
;             STREET: 666 Fifth Avenue
;             CITY: New York City
;             STATE: New York
;             COUNTRY: USA
;             ZIP: 10103
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.25 inch, 1.44mb
;             COMPUTER: IBM PS/2
;             OPERATING SYSTEM: PC-DOS
;             SOFTWARE: Wordperfect
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/127,427
;             FILING DATE: 23-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/333,527
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 60/089,322
;             FILING DATE: June 15, 1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mary Anne Schofield
;             REGISTRATION NUMBER: 36,669
;             REFERENCE/DOCKET NUMBER: KL/JIC 202.1 - JEL
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (212) 318-3000
;             TELEFAX: (212) 752-5958
;                      582015
;   INFORMATION FOR SEQ ID NO: 7:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 57
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-127-427-7

  Query Match             0.8%;  Score 19.6;  DB 9;  Length 57;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+05;
  Matches   28;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   2110 TTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGAT 2151
          ||||||||  | || | || | | | | | ||    ||||| || |
Db     50 TTTTTTTTTTTTTATAAATCTCTCTATAAATCTGATTTGTGTT 9


RESULT 44
US-10-146-502-2169/c
; Sequence 2169, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527C2
;  CURRENT APPLICATION NUMBER: US/10/146,502
;  CURRENT FILING DATE:  2002-05-14
;  NUMBER OF SEQ ID NOS: 2241
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2169
;   LENGTH: 57
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 13
;   OTHER INFORMATION: n = A,T,C or G
US-10-146-502-2169

  Query Match             0.8%;  Score 19.6;  DB 9;  Length 57;
  Best Local Similarity   65.1%;  Pred. No. 1.2e+05;
  Matches   28;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   1760 AAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
          ||||||  ||  ||  | | |||||||||   |||  |||  |
Db     54 AAAAAAAAAAAAAAAAGGGAAAAAAAAAAAAAAAAAAAAAGNA 12


RESULT 45
US-09-333-527-7/c
; Sequence 7, Application US/09333527
; Patent No. US20020078472A1
;  GENERAL INFORMATION:
;    APPLICANT:  Paul CHRISTOU; Eva STROGER; Rainer FISCHER; Carmen MARTIN-VAQUERO; Stefa
;    TITLE OF INVENTION:  METHODS AND MEANS FOR EXPRESSION OF MAMMALIAN POLYPEPTIDES
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski L.L.P.
;      STREET:  666 Fifth Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.25 inch, 1.44mb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/333,527
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/089,322
;      FILING DATE:  June 15, 1998
;    ATTORNEY/AGENT INFORMATION:
```

```
;       NAME:  Mary Anne Schofield
;       REGISTRATION NUMBER:  36,669
;       REFERENCE/DOCKET NUMBER:  KL/JIC 202.1 - JEL
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 318-3000
;       TELEFAX:  (212) 752-5958
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  57
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-09-333-527-7

  Query Match             0.8%;  Score 19.6;  DB 10;  Length 57;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+05;
  Matches   28;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   2110 TTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGAT 2151
          |||||||| | || | || | | | | ||   ||||| || |
Db     50 TTTTTTTTTTTTATAAATCTCTCTATAAATCTGATTTGTGTT 9


RESULT 46
US-10-085-906-66
; Sequence 66, Application US/10085906
; Publication No. US20030054371A1
; GENERAL INFORMATION:
;  APPLICANT: Ying, Vincent
;  APPLICANT:  Wu, Paul
;  APPLICANT:  Gray, Gary S.
;  TITLE OF INVENTION: POLYMORPHIC ELEMENTS IN THE
;  TITLE OF INVENTION:  COSTIMULATORY RECEPTOR LOCUS AND USES THEREOF
;  FILE REFERENCE: GNN-5343CP2
;  CURRENT APPLICATION NUMBER: US/10/085,906
;  CURRENT FILING DATE:  2002-02-27
;  PRIOR APPLICATION NUMBER: US 60/126,215
;  PRIOR FILING DATE: 1999-03-25
;  PRIOR APPLICATION NUMBER: US 09/534,061
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: PCT/US00/07938
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 545
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 66
;   LENGTH: 54
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-085-906-66

  Query Match             0.8%;  Score 19.4;  DB 9;  Length 54;
  Best Local Similarity   64.4%;  Pred. No. 1.3e+05;
  Matches   29;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
          | |||||| |||||  || ||| | |||   ||  |  || |||
Db      2 TATTTTTTGTATATTTATTTATTTATTTATTTTAATTAATTAATT 46


RESULT 47
US-09-983-965-452/c
; Sequence 452, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 37-21(10297)C
;  CURRENT APPLICATION NUMBER: US/09/983,965
;  CURRENT FILING DATE:  2001-10-26
;  PRIOR APPLICATION NUMBER: US 09/465,231
;  PRIOR FILING DATE: 1999-12-15
;  PRIOR APPLICATION NUMBER: US 60/113,678
;  PRIOR FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 452
;   LENGTH: 54
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 53-BOVMS1-015-Q1-E1-F10
US-09-983-965-452

  Query Match             0.8%;  Score 19.4;  DB 10;  Length 54;
  Best Local Similarity   64.4%;  Pred. No. 1.3e+05;
  Matches   29;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
          | |||||||| | |  || | | ||||||    |   |||||  ||
Db     47 TTTTTTTTTTTTTTTTATTTTTTTTTTTCTTTTTTTTTTTTTTT 3


RESULT 48
US-09-263-959-519
; Sequence 519, Application US/09263959
; Patent No. US20020150891A1
;  GENERAL INFORMATION:
;    APPLICANT:  Hood, Leroy E.
;    APPLICANT:  Rowen, Lee
;    APPLICANT:  Koop, Ben F.
;    TITLE OF INVENTION:  DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH UTI
;    NUMBER OF SEQUENCES:  1279
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/263,959
;      FILING DATE:  05-MAR-1999
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  920010.426C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  519:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  54 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-263-959-519

  Query Match             0.8%;  Score 19.4;  DB 10;  Length 54;
  Best Local Similarity   64.4%;  Pred. No. 1.3e+05;
  Matches   29;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   1762 AAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATAGTAT 1806
          ||||| || |||       | |||| ||  ||||  ||||||  ||
Db      1 AAAATAAAACAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 45


RESULT 49
```

```
US-10-046-935-335/c
; Sequence 335, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527C1
;  CURRENT APPLICATION NUMBER: US/10/046,935
;  CURRENT FILING DATE:  2002-01-15
;  NUMBER OF SEQ ID NOS: 2239
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 335
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 10, 13
;   OTHER INFORMATION: n = A,T,C or G
US-10-046-935-335

  Query Match             0.8%;  Score 19.4;  DB 9;  Length 60;
  Best Local Similarity   58.2%;  Pred. No. 1.4e+05;
  Matches   32;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy   1756 TTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATAGTATTGCT 1810
          ||||||  ||  ||  ||      | |||||||   |||  |||| |  |  |||
Db     60 TTTTAACCAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAANAANAGCT 6


RESULT 50
US-09-878-178-335/c
; Sequence 335, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527
;  CURRENT APPLICATION NUMBER: US/09/878,178
;  CURRENT FILING DATE:  2001-06-08
;  NUMBER OF SEQ ID NOS: 2237
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 335
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(60)
;   OTHER INFORMATION: n = A,T,C or G
US-09-878-178-335

  Query Match             0.8%;  Score 19.4;  DB 9;  Length 60;
  Best Local Similarity   58.2%;  Pred. No. 1.4e+05;
  Matches   32;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy   1756 TTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATAGTATTGCT 1810
          ||||||  ||  ||  ||      | |||||||   |||  |||| |  |  |||
Db     60 TTTTAACCAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAANAANAGCT 6


RESULT 51
US-10-208-155-6
; Sequence 6, Application US/10208155
; Publication No. US20030013171A1
;  GENERAL INFORMATION:
;       APPLICANT: Yang et al.
;       TITLE OF INVENTION: BCL-X{SYMBOL 103  \f "Symbol"}, A NOVEL BCL-X
;                           ISOFORM, AND USES RELATED THERETO
;       NUMBER OF SEQUENCES: 23
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: LAHIVE & COCKFIELD
;            STREET: 60 State Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02109-1875
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/208,155
;            FILING DATE: 29-Jul-2002
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/899,367
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Amy E. Mandragouras
;            REGISTRATION NUMBER: 36,207
;            REFERENCE/DOCKET NUMBER: DFN-019
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)227-7400
;            TELEFAX: (617)227-5941
;  INFORMATION FOR SEQ ID NO: 6:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 60 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-208-155-6

  Query Match             0.8%;  Score 19.4;  DB 9;  Length 60;
  Best Local Similarity   70.3%;  Pred. No. 1.4e+05;
  Matches   26;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy    150 GGATGAGCTGGAGGAGGGGGGTCAGTTAAATGAAAGC 186
          || |||  ||||| |||  ||||| | | |||| ||
Db      2 GGGTGATGTGGAGCTGGGATGTCAGGTCACTGAATGC 38


RESULT 52
US-10-146-502-335/c
; Sequence 335, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527C2
;  CURRENT APPLICATION NUMBER: US/10/146,502
;  CURRENT FILING DATE:  2002-05-14
;  NUMBER OF SEQ ID NOS: 2241
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 335
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 10, 13
```

```
;   OTHER INFORMATION: n = A,T,C or G
US-10-146-502-335

  Query Match             0.8%;  Score 19.4;  DB 9;  Length 60;
  Best Local Similarity   58.2%;  Pred. No. 1.4e+05;
  Matches   32;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy   1756 TTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATAGTATTGCT 1810
          ||||||  ||  ||  ||       | |||||||   |||  |||| |  |  |||
Db     60 TTTTAACCAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAANAANAGCT 6


RESULT 53
US-09-730-289B-2736
; Sequence 2736, Application US/09730289B
; Publication No. US20030050259A1
; GENERAL INFORMATION:
;  APPLICANT: Ribozyme Pharmaceuticals, Inc.
;  APPLICANT:  Blatt, Larry
;  APPLICANT:  McSwiggen, Jim
;  TITLE OF INVENTION: Method and Reagent for Treatment of Cardiac Disease
;  FILE REFERENCE: MBHB00-864-A (400/006)
;  CURRENT APPLICATION NUMBER: US/09/730,289B
;  CURRENT FILING DATE:  2000-12-05
;  PRIOR APPLICATION NUMBER: US 60/169,100
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 3897
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2736
;   LENGTH: 37
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Enzymatic Nucleic Acid
US-09-730-289B-2736

  Query Match             0.8%;  Score 19.2;  DB 9;  Length 37;
  Best Local Similarity   56.2%;  Pred. No. 1.2e+05;
  Matches   18;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    176 TAAATGAAAGCATGGACCATGGGGGAGTTGGA 207
          : ||:||  |||:| || |:| | |||::| |
Db      5 UCAAUGAUGGCAUGCACUAUGCGCGAGUUGCA 36


RESULT 54
US-09-263-959-213/c
; Sequence 213, Application US/09263959
; Patent No. US20020150891A1
;  GENERAL INFORMATION:
;    APPLICANT:  Hood, Leroy E.
;    APPLICANT:  Rowen, Lee
;    APPLICANT:  Koop, Ben F.
;    TITLE OF INVENTION:  DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH UTI
;    NUMBER OF SEQUENCES:  1279
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/263,959
;      FILING DATE:  05-MAR-1999
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  920010.426C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  213:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  39 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-263-959-213

  Query Match             0.8%;  Score 19.2;  DB 10;  Length 39;
  Best Local Similarity   87.5%;  Pred. No. 1.2e+05;
  Matches   21;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   1911 AGGATGTTGCTCTGAATCACTGTG 1934
          ||||||| |||||| |||| ||||
Db     24 AGGATGTGGCTCTGCATCAATGTG 1


RESULT 55
US-09-753-436-101/c
; Sequence 101, Application US/09753436
; Patent No. US20010029293A1
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Vazeux, Rosemay
;    TITLE OF INVENTION:  ICAM-Related Materials and Methods
;    NUMBER OF SEQUENCES:  120
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/753,436
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/382,289
;      FILING DATE:
;      APPLICATION NUMBER:  US 08/487,113
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,754
;      FILING DATE:  05-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/102,852
;      FILING DATE:  05-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/009,266
;      FILING DATE:  22-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/894,061
;      FILING DATE:  05-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/889,724
;      FILING DATE:  26-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/827,689
;      FILING DATE:  27-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams, Joseph A., Jr.
;      REGISTRATION NUMBER:  38,659
```

```
;      REFERENCE/DOCKET NUMBER:  33282
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  101:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  43 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-09-753-436-101

  Query Match             0.8%;  Score 19.2;  DB 10;  Length 43;
  Best Local Similarity   67.5%;  Pred. No. 1.3e+05;
  Matches   27;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    189 GGACCATGGGGGAGTTGGACCATATGAACTTGGCATGGAA 228
          |||||| |  |||| || | | || | |||| | ||  ||
Db     40 GGACCAAGCTGGAGCTGAAACGTAAGTACTTAGGATCCAA 1


RESULT 56
US-09-263-959-148/c
; Sequence 148, Application US/09263959
; Patent No. US20020150891A1
;  GENERAL INFORMATION:
;    APPLICANT:  Hood, Leroy E.
;    APPLICANT:  Rowen, Lee
;    APPLICANT:  Koop, Ben F.
;    TITLE OF INVENTION:  DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH UTI
;    NUMBER OF SEQUENCES:  1279
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/263,959
;      FILING DATE:  05-MAR-1999
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  920010.426C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  148:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-263-959-148

  Query Match             0.8%;  Score 19.2;  DB 10;  Length 46;
  Best Local Similarity   87.5%;  Pred. No. 1.4e+05;
  Matches   21;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   1911 AGGATGTTGCTCTGAATCACTGTG 1934
          ||||||| |||||| |||| ||||
Db     24 AGGATGTGGCTCTGCATCAATGTG 1


RESULT 57
US-09-999-724-14
; Sequence 14, Application US/09999724
; Publication No. US20030022355A1
; GENERAL INFORMATION:
;  APPLICANT: WICKHAM, THOMAS J.
;  APPLICANT:  KOVESDI, IMRE
;  APPLICANT:  BROUGH, DOUGLAS E.
;  TITLE OF INVENTION: VECTORS AND METHODS FOR GENE TRANSFER
;  FILE REFERENCE: 212960
;  CURRENT APPLICATION NUMBER: US/09/999,724
;  CURRENT FILING DATE:  2001-10-24
;  PRIOR APPLICATION NUMBER: US 09/101,751
;  PRIOR FILING DATE: 1999-01-29
;  PRIOR APPLICATION NUMBER: WO 96US19150
;  PRIOR FILING DATE: 1996-11-27
;  PRIOR APPLICATION NUMBER: US 08/700,846
;  PRIOR FILING DATE: 1996-08-21
;  PRIOR APPLICATION NUMBER: US 08/701,124
;  PRIOR FILING DATE: 1996-08-21
;  PRIOR APPLICATION NUMBER: US 08/563,368
;  PRIOR FILING DATE: 1995-11-28
;  NUMBER OF SEQ ID NOS: 94
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 48
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
US-09-999-724-14

  Query Match             0.8%;  Score 19.2;  DB 9;  Length 48;
  Best Local Similarity   60.0%;  Pred. No. 1.4e+05;
  Matches   24;  Conservative    3;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   1761 AAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAA 1800
          || ||  ||:|  : |:|  |||||||.  |||  ||||
Db      8 AAUAAAGAAUCGUUUGUGUUAAAAAAAAAAAAAAAAAAAA 47


RESULT 58
US-09-796-679-18/c
; Sequence 18, Application US/09796679
; Publication No. US20030013076A1
; GENERAL INFORMATION:
;  APPLICANT: Robinson, Anthony J
;  APPLICANT:  Lyttle, David J
;  APPLICANT:  University of Otago
;  TITLE OF INVENTION: Parapoxvirus vectors
;  FILE REFERENCE: 23607 MRB
;  CURRENT APPLICATION NUMBER: US/09/796,679
;  CURRENT FILING DATE:  2001-03-30
;  PRIOR APPLICATION NUMBER: 09/155,421
;  PRIOR FILING DATE: 1998-09-29
;  PRIOR APPLICATION NUMBER: PCT/NZ97/00040
;  PRIOR FILING DATE: 1997-03-27
;  PRIOR APPLICATION NUMBER: NZ 286284
;  PRIOR FILING DATE: 1996-03-29
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
;   LENGTH: 49
;   TYPE: DNA
;   ORGANISM: Orf virus strain NZ-2
US-09-796-679-18

  Query Match             0.8%;  Score 19.2;  DB 9;  Length 49;
  Best Local Similarity   67.5%;  Pred. No. 1.4e+05;
  Matches   27;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   2177 AACTTTAAAATGCGAAATTATTGGTTGGTGTGAAGAAAGC 2216
          | ||||  |  |||  |||||||  |||||  | |  ||   ||
```

```
Db      48 ATCTTTTATGTGCAGAATTATTCGTTGGCGCGGAGCTCGC 9


RESULT 59
US-10-085-906-204/c
; Sequence 204, Application US/10085906
; Publication No. US20030054371A1
; GENERAL INFORMATION:
;  APPLICANT: Ying, Vincent
;  APPLICANT:  Wu, Paul
;  APPLICANT:  Gray, Gary S.
;  TITLE OF INVENTION: POLYMORPHIC ELEMENTS IN THE
;  TITLE OF INVENTION:  COSTIMULATORY RECEPTOR LOCUS AND USES THEREOF
;  FILE REFERENCE: GNN-5343CP2
;  CURRENT APPLICATION NUMBER: US/10/085,906
;  CURRENT FILING DATE:  2002-02-27
;  PRIOR APPLICATION NUMBER: US 60/126,215
;  PRIOR FILING DATE: 1999-03-25
;  PRIOR APPLICATION NUMBER: US 09/534,061
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: PCT/US00/07938
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 545
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 204
;   LENGTH: 49
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-085-906-204

  Query Match             0.8%;  Score 19.2;  DB 9;  Length 49;
  Best Local Similarity   67.5%;  Pred. No. 1.4e+05;
  Matches   27;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy     911 GGGAGAATAGAAAGAAAACAATTGACAAAATCCTCAAATG 950
           | ||| | ||| |||||| ||   | |||| | | ||| |
Db      41 GAGAGGAAAGAGAGAAAAAAAGATAGAAAAACATAAAAAG 2


RESULT 60
US-09-907-900-52/c
; Sequence 52, Application US/09907900
; Patent No. US20020172997A1
; GENERAL INFORMATION:
;  APPLICANT: Hartley, James L.
;  APPLICANT:  Brasch, Michael A.
;  APPLICANT:  Temple, Gary F.
;  APPLICANT:  Fox, Donna K.
;  TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having
;  TITLE OF INVENTION:  Recombination Sites
;  FILE REFERENCE: 0942.2850004
;  CURRENT APPLICATION NUMBER: US/09/907,900
;  CURRENT FILING DATE:  2001-07-19
;  PRIOR APPLICATION NUMBER: 09/177,387
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
;   LENGTH: 53
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: oligonucleotide
US-09-907-900-52

  Query Match             0.8%;  Score 19.2;  DB 9;  Length 53;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+05;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    2009 AATTAGCATGCAAGCTTGGTCAAACTTTTCCC 2040
           ||||||| |||    |||  ||||||| ||||
Db      33 AATTAGCCTGCTTTTTTGTACAAACTTGTCCC 2


RESULT 61
US-09-907-719-52/c
; Sequence 52, Application US/09907719
; Publication No. US20020192819A1
; GENERAL INFORMATION:
;  APPLICANT: Hartley, James L.
;  APPLICANT:  Brasch, Michael A.
;  APPLICANT:  Temple, Gary F.
;  APPLICANT:  Fox, Donna K.
;  TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having
;  TITLE OF INVENTION:  Recombination Sites
;  FILE REFERENCE: 0942.2850004
;  CURRENT APPLICATION NUMBER: US/09/907,719
;  CURRENT FILING DATE:  2001-07-19
;  PRIOR APPLICATION NUMBER: US/09/177,387
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
;   LENGTH: 53
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: oligonucleotide
US-09-907-719-52

  Query Match             0.8%;  Score 19.2;  DB 9;  Length 53;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+05;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    2009 AATTAGCATGCAAGCTTGGTCAAACTTTTCCC 2040
           ||||||| |||    |||  ||||||| ||||
Db      33 AATTAGCCTGCTTTTTTGTACAAACTTGTCCC 2


RESULT 62
US-09-855-797A-52/c
; Sequence 52, Application US/09855797A
; Patent No. US20020094574A1
; GENERAL INFORMATION:
;  APPLICANT: Hartley, James L.
;  APPLICANT:  Brasch, Michael A.
;  APPLICANT:  Temple, Gary F.
;  APPLICANT:  Fox, Donna K.
;  TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having
;  TITLE OF INVENTION:  Recombination Sites
;  FILE REFERENCE: 0942.2850008
;  CURRENT APPLICATION NUMBER: US/09/855,797A
;  CURRENT FILING DATE:  2001-05-16
;  PRIOR APPLICATION NUMBER: 09/296,281
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: US 60/065,930
;  PRIOR FILING DATE: 1997-10-24
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
;   LENGTH: 53
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: oligonucleotide
US-09-855-797A-52

  Query Match             0.8%;  Score 19.2;  DB 10;  Length 53;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+05;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    2009 AATTAGCATGCAAGCTTGGTCAAACTTTTCCC 2040
           ||||||| |||    |||  ||||||| ||||
```

```
Db     33 AATTAGCCTGCTTTTTTGTACAAACTTGTCCC 2


RESULT 63
US-09-970-308-8
; Sequence 8, Application US/09970308
; Patent No. US20020045193A1
; GENERAL INFORMATION:
;  APPLICANT: BRIZZARD, BILLY L.
;  APPLICANT:  HERNAN, RON
;  TITLE OF INVENTION: PURIFICATION OF RECOMBINANT PROTEINS FUSED TO MULTIPLE
;  TITLE OF INVENTION:  EPITOPES
;  FILE REFERENCE: SGM 6933.2
;  CURRENT APPLICATION NUMBER: US/09/970,308
;  CURRENT FILING DATE:  2001-10-03
;  PRIOR APPLICATION NUMBER: 09/415,000
;  PRIOR FILING DATE: 1999-10-08
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 54
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Primer
US-09-970-308-8

  Query Match             0.8%;  Score 19.2;  DB 10;  Length 54;
  Best Local Similarity   67.5%;  Pred. No. 1.5e+05;
  Matches   27;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    377 GGAAAATATTTGCCATGAAGGTGCTTAAAAAGGCAATGAT 416
          |||  | |    |||||| |||| ||| ||||   |||||
Db     15 GGACTACAAAGACCATGACGGTGATTATAAAGATCATGAT 54


RESULT 64
US-09-970-308-9/c
; Sequence 9, Application US/09970308
; Patent No. US20020045193A1
; GENERAL INFORMATION:
;  APPLICANT: BRIZZARD, BILLY L.
;  APPLICANT:  HERNAN, RON
;  TITLE OF INVENTION: PURIFICATION OF RECOMBINANT PROTEINS FUSED TO MULTIPLE
;  TITLE OF INVENTION:  EPITOPES
;  FILE REFERENCE: SGM 6933.2
;  CURRENT APPLICATION NUMBER: US/09/970,308
;  CURRENT FILING DATE:  2001-10-03
;  PRIOR APPLICATION NUMBER: 09/415,000
;  PRIOR FILING DATE: 1999-10-08
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 54
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Primer
US-09-970-308-9

  Query Match             0.8%;  Score 19.2;  DB 10;  Length 54;
  Best Local Similarity   67.5%;  Pred. No. 1.5e+05;
  Matches   27;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    377 GGAAAATATTTGCCATGAAGGTGCTTAAAAAGGCAATGAT 416
          |||  | |    |||||| |||| ||| ||||   |||||
Db     40 GGACTACAAAGACCATGACGGTGATTATAAAGATCATGAT 1


RESULT 65
US-09-426-548-41
; Sequence 41, Application US/09426548
; Patent No. US20010044936A1
; GENERAL INFORMATION:
;  APPLICANT: Robbins, David
;  APPLICANT: Lin-Goerke, Juili L.
;  APPLICANT: Ling, Jessica
;  TITLE OF INVENTION: No. US20010044936A1el Mutations in Human MLH1 and MSH2 Genes Usefu
;  TITLE OF INVENTION: Diagnosing Colorectal Cancer
;  FILE REFERENCE: DEX-0054
;  CURRENT APPLICATION NUMBER: US/09/426,548
;  CURRENT FILING DATE: 1999-10-22
;  NUMBER OF SEQ ID NOS: 192
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
;   LENGTH: 57
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-426-548-41

  Query Match             0.8%;  Score 19.2;  DB 10;  Length 57;
  Best Local Similarity   67.5%;  Pred. No. 1.6e+05;
  Matches   27;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   2117 TATATAAATATATATTTTTCAAATAGATTTTTGATTCAGC 2156
          |||||||| | || |||| ||  ||   | ||| || | |
Db     15 TATATAAAAAAATCTTTTACATTTATTATCTTGGTTTATC 54


RESULT 66
US-09-902-941-2002
; Sequence 2002, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Johnson, Jeffrey C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Marnerakis, Margarita
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  McNabb, Andria
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C17
;  CURRENT APPLICATION NUMBER: US/09/902,941
;  CURRENT FILING DATE:  2001-07-10
;  NUMBER OF SEQ ID NOS: 2002
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2002
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-902-941-2002

  Query Match             0.8%;  Score 19.2;  DB 9;  Length 60;
  Best Local Similarity   58.9%;  Pred. No. 1.6e+05;
  Matches   33;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy    652 TTACATCAAAAGGGGATCATCTACAGAGACCTGAAGCCGGAGAATATCATGCTTAA 707
          | | | | |||| ||| ||     | |  || ||||| || ||| | ||  |  ||
Db      2 TAATACCTAAAGAGGAACACTGTAAAATGCCAGAAGCAGGTGAAGAGCAACCACAA 57


RESULT 67
US-10-017-754-2002
; Sequence 2002, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Johnson, Jeffrey C.
```

```
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Marnerakis, Margarita
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  McNabb, Andria
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C18
;  CURRENT APPLICATION NUMBER: US/10/017,754
;  CURRENT FILING DATE:  2001-10-29
;  NUMBER OF SEQ ID NOS: 2004
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2002
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-017-754-2002

  Query Match             0.8%;  Score 19.2;  DB 9;  Length 60;
  Best Local Similarity   58.9%;  Pred. No. 1.6e+05;
  Matches   33;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy    652 TTACATCAAAAGGGGATCATCTACAGAGACCTGAAGCCGGAGAATATCATGCTTAA 707
          | | | | |||| ||| ||     | |  || ||||| || ||| | ||  |  ||
Db      2 TAATACCTAAAGAGGAACACTGTAAAATGCCAGAAGCAGGTGAAGAGCAACCACAA 57


RESULT 68
US-10-085-906-255/c
; Sequence 255, Application US/10085906
; Publication No. US20030054371A1
; GENERAL INFORMATION:
;  APPLICANT: Ying, Vincent
;  APPLICANT:  Wu, Paul
;  APPLICANT:  Gray, Gary S.
;  TITLE OF INVENTION: POLYMORPHIC ELEMENTS IN THE
;  TITLE OF INVENTION:  COSTIMULATORY RECEPTOR LOCUS AND USES THEREOF
;  FILE REFERENCE: GNN-5343CP2
;  CURRENT APPLICATION NUMBER: US/10/085,906
;  CURRENT FILING DATE:  2002-02-27
;  PRIOR APPLICATION NUMBER: US 60/126,215
;  PRIOR FILING DATE: 1999-03-25
;  PRIOR APPLICATION NUMBER: US 09/534,061
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: PCT/US00/07938
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 545
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 255
;   LENGTH: 46
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-085-906-255

  Query Match             0.8%;  Score 19;  DB 9;  Length 46;
  Best Local Similarity   65.1%;  Pred. No. 1.5e+05;
  Matches   28;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2110 TTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
          ||| ||||||  | |  | ||| || ||  | |  |||||   ||
Db     43 TTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1


RESULT 69
US-09-827-289-20
; Sequence 20, Application US/09827289
; Patent No. US20020009716A1
; GENERAL INFORMATION:
;  APPLICANT: Abarzua, Patricio
;  TITLE OF INVENTION: Process for Allele Discrimination Using Primer
;  TITLE OF INVENTION:  Extension
;  FILE REFERENCE: 469290-55
;  CURRENT APPLICATION NUMBER: US/09/827,289
;  CURRENT FILING DATE:  2001-04-05
;  PRIOR APPLICATION NUMBER: U.S. 60/194843
;  PRIOR FILING DATE: 2000-04-05
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
;   LENGTH: 46
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: P1 primer for
;   OTHER INFORMATION: use in allele discrimination
US-09-827-289-20

  Query Match             0.8%;  Score 19;  DB 10;  Length 46;
  Best Local Similarity   65.1%;  Pred. No. 1.5e+05;
  Matches   28;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2111 TTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTC 2153
          ||||||| | |   | | | ||||| ||| |   |||  ||||
Db      1 TTTTTTTTTTTTTTTTTTTTTTTTTTAAAAATCATTTCTATTC 43


RESULT 70
US-09-827-289-24
; Sequence 24, Application US/09827289
; Patent No. US20020009716A1
; GENERAL INFORMATION:
;  APPLICANT: Abarzua, Patricio
;  TITLE OF INVENTION: Process for Allele Discrimination Using Primer
;  TITLE OF INVENTION:  Extension
;  FILE REFERENCE: 469290-55
;  CURRENT APPLICATION NUMBER: US/09/827,289
;  CURRENT FILING DATE:  2001-04-05
;  PRIOR APPLICATION NUMBER: U.S. 60/194843
;  PRIOR FILING DATE: 2000-04-05
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
;   LENGTH: 46
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: P1 primer for
;   OTHER INFORMATION: use in allele discrimination
US-09-827-289-24

  Query Match             0.8%;  Score 19;  DB 10;  Length 46;
  Best Local Similarity   65.1%;  Pred. No. 1.5e+05;
  Matches   28;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2111 TTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTC 2153
          ||||||| | |   | | | ||||| ||| |   |||  ||||
Db      1 TTTTTTTTTTTTTTTTTTTTTTTTTTAAAAATCATTTCTATTC 43


RESULT 71
US-09-938-842A-3904
; Sequence 3904, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
```

```
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 3904
;   LENGTH: 53
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-3904

  Query Match             0.8%;  Score 19;  DB 9;  Length 53;
  Best Local Similarity   71.4%;  Pred. No. 1.7e+05;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   2085 GTTTTACGTGCAAACAACCTGAATCTTTTTTTTAT 2119
          || ||||    | | ||| ||||| |||||||| |
Db      1 GTATTACTAAAAGAAAACTTGAATGTTTTTTTTTT 35


RESULT 72
US-09-263-959-519/c
; Sequence 519, Application US/09263959
; Patent No. US20020150891A1
;  GENERAL INFORMATION:
;    APPLICANT:  Hood, Leroy E.
;    APPLICANT:  Rowen, Lee
;    APPLICANT:  Koop, Ben F.
;    TITLE OF INVENTION:  DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH UTI
;    NUMBER OF SEQUENCES:  1279
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/263,959
;      FILING DATE:  05-MAR-1999
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  920010.426C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  519:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  54 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-263-959-519

  Query Match             0.8%;  Score 19;  DB 10;  Length 54;
  Best Local Similarity   65.1%;  Pred. No. 1.7e+05;
  Matches   28;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2110 TTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
          ||| |||||| | | | ||| || ||  | |  |||||   ||
Db     53 TTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 11


RESULT 73
US-10-085-906-18/c
; Sequence 18, Application US/10085906
; Publication No. US20030054371A1
; GENERAL INFORMATION:
;  APPLICANT: Ying, Vincent
;  APPLICANT:  Wu, Paul
;  APPLICANT:  Gray, Gary S.
;  TITLE OF INVENTION: POLYMORPHIC ELEMENTS IN THE
;  TITLE OF INVENTION:  COSTIMULATORY RECEPTOR LOCUS AND USES THEREOF
;  FILE REFERENCE: GNN-5343CP2
;  CURRENT APPLICATION NUMBER: US/10/085,906
;  CURRENT FILING DATE:  2002-02-27
;  PRIOR APPLICATION NUMBER: US 60/126,215
;  PRIOR FILING DATE: 1999-03-25
;  PRIOR APPLICATION NUMBER: US 09/534,061
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: PCT/US00/07938
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 545
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
;   LENGTH: 55
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-085-906-18

  Query Match             0.8%;  Score 19;  DB 9;  Length 55;
  Best Local Similarity   81.5%;  Pred. No. 1.7e+05;
  Matches   22;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy   2045 AAAATGGGAAGGCAAAGACAAAAGAAA 2071
          ||||  |||| | ||||| ||||||||
Db     38 AAAAAAGGAAAGAAAAGAAAAAAGAAA 12


RESULT 74
US-10-007-132-59/c
; Sequence 59, Application US/10007132
; Publication No. US20030027254A1
;   GENERAL INFORMATION:
;        APPLICANT: Bard, Jonathan A
;                   Borowsky, Beth
;                   Smith, Kelli E
;        TITLE OF INVENTION: DNA ENCODING GALANIN GALR3 RECEPTORS
;                            AND USES THEREOF
;        NUMBER OF SEQUENCES: 65
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Cooper & Dunham LLP
;             STREET: 1185 Avenue of the Americas
;             CITY: New York
;             STATE: New York
;             COUNTRY: U.S.A.
;             ZIP: 10036
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/007,132
;             FILING DATE: 03-Dec-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/058,333
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: White, John P
;             REGISTRATION NUMBER: 28,678
;             REFERENCE/DOCKET NUMBER: 52241-E/JPW/KDB
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 212 278 0400
;             TELEFAX: 212 391 0525
;   INFORMATION FOR SEQ ID NO: 59:
;        SEQUENCE CHARACTERISTICS:
```

```
;            LENGTH: 57 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;        HYPOTHETICAL: NO
;        ANTI-SENSE: NO
;        SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-10-007-132-59

  Query Match             0.8%;  Score 19;  DB 9;  Length 57;
  Best Local Similarity   71.4%;  Pred. No. 1.8e+05;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1768 AATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
          ||| ||   ||||||||||||   |||  |||| |
Db     47 AATAAACCCTGCAAAAAAAAAAAAAAAAAAAAAAAA 13


RESULT 75
US-09-944-036-31/c
; Sequence 31, Application US/09944036
; Patent No. US20020055095A1
; GENERAL INFORMATION:
;  APPLICANT: YANG, Yeasing Y.
;  APPLICANT:  BRENTANO, Steven T.
;  APPLICANT:  BABOLA, Odile
;  APPLICANT:  TRAN, Nathalie
;  APPLICANT:  VERNET, Guy
;  TITLE OF INVENTION: AMPLIFICATION OF HIV-1 SEQUENCES FOR DETECTION OF
;  TITLE OF INVENTION:  SEQUENCES ASSOCIATED WITH DRUG-RESISTANCE MUTATIONS
;  FILE REFERENCE: GP114-02.UT
;  CURRENT APPLICATION NUMBER: US/09/944,036
;  CURRENT FILING DATE:  2001-08-31
;  PRIOR APPLICATION NUMBER: US 60/229,790
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
;   LENGTH: 57
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Capture
;   OTHER INFORMATION: oligonucleotide
US-09-944-036-31

  Query Match             0.8%;  Score 19;  DB 10;  Length 57;
  Best Local Similarity   65.1%;  Pred. No. 1.8e+05;
  Matches   28;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGA 2150
          | ||||||||  | |   | | | ||||| |||  |.||| | |
Db     56 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAACGGGTTTATTA 14


RESULT 76
US-09-943-286-7/c
; Sequence 7, Application US/09943286
; Patent No. US20020106668A1
; GENERAL INFORMATION:
;  APPLICANT: Nunomura, Kiyotada
;  TITLE OF INVENTION: POLYNUCLEOTIDE AMPLIFICATION METHOD
;  FILE REFERENCE: GP104-02.UT
;  CURRENT APPLICATION NUMBER: US/09/943,286
;  CURRENT FILING DATE:  2001-08-30
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
;   LENGTH: 57
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Sequence of the (-)4258 A30 capture probe.
US-09-943-286-7

  Query Match             0.8%;  Score 19;  DB 10;  Length 57;
  Best Local Similarity   65.1%;  Pred. No. 1.8e+05;
  Matches   28;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGA 2150
          | ||||||||  | |   | | | ||||| |||  | ||| | |
Db     56 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAACGGGTTTATTA 14


RESULT 77
US-09-954-692-17
; Sequence 17, Application US/09954692
; Publication No. US20030027156A1
; GENERAL INFORMATION:
;  APPLICANT: Patten, Phillip
;  APPLICANT:  Stemmer, Willem P.C.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR POLYPEPTIDE ENGINEERING
;  FILE REFERENCE: 02-020500US
;  CURRENT APPLICATION NUMBER: US/09/954,692
;  CURRENT FILING DATE:  2001-09-12
;  PRIOR APPLICATION NUMBER: US/08/769,062
;  PRIOR FILING DATE: 1996-12-18
;  PRIOR APPLICATION NUMBER: 08/425,684
;  PRIOR FILING DATE: 1995-04-18
;  PRIOR APPLICATION NUMBER: 08/537,874
;  PRIOR FILING DATE: 1995-10-30
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: degenerate
;   OTHER INFORMATION: oligonucleotide used for codon usage library
US-09-954-692-17

  Query Match             0.8%;  Score 19;  DB 9;  Length 60;
  Best Local Similarity   57.1%;  Pred. No. 1.8e+05;
  Matches   28;  Conservative    4;  Mismatches   17;  Indels    0;  Gaps    0;

Qy    144 CTCTGAGGATGAGCTGGAGGAGGGGGGTCAGTTAAATGAAAGCATGGAC 192
          |||||| | ||| |||  :| :|  | :||: | || |   ||  ||||
Db      4 CTCTGACGCTGACCTGCCDGCDGAYGCDCARATGAACGGTTGCCAGGAC 52


RESULT 78
US-09-559-671A-17
; Sequence 17, Application US/09559671A
; Patent No. US20020051976A1
; GENERAL INFORMATION:
;  APPLICANT: Patten, Phillip
;  APPLICANT:  Stemmer, Willem P.C.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR POLYPEPTIDE ENGINEERING
;  FILE REFERENCE: 02-020503US
;  CURRENT APPLICATION NUMBER: US/09/559,671A
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: 08/769,062
;  PRIOR FILING DATE: 1996-12-18
;  PRIOR APPLICATION NUMBER: 08/198,431
;  PRIOR FILING DATE: 1994-02-17
;  PRIOR APPLICATION NUMBER: 08/425,684
;  PRIOR FILING DATE: 1995-04-18
;  PRIOR APPLICATION NUMBER: 08/537,874
;  PRIOR FILING DATE: 1995-10-30
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
```

```
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: degenerate
;   OTHER INFORMATION: oligonucleotide used for codon usage library
US-09-559-671A-17

  Query Match             0.8%;  Score 19;  DB 10;  Length 60;
  Best Local Similarity   57.1%;  Pred. No. 1.8e+05;
  Matches   28;  Conservative    4;  Mismatches   17;  Indels    0;  Gaps    0;

Qy    144 CTCTGAGGATGAGCTGGAGGAGGGGGGTCAGTTAAATGAAAGCATGGAC 192
          |||||| | ||| |||  :| :|  | :||: | || |   ||  ||||
Db      4 CTCTGACGCTGACCTGCCDGCDGAYGCDCARATGAACGGTTGCCAGGAC 52


RESULT 79
US-09-877-478-4121
; Sequence 4121, Application US/09877478
; Publication No. US20030068301A1
; GENERAL INFORMATION:
;  APPLICANT: Ribozyme Pharmaceuticals, Inc.
;  APPLICANT:  Draper, Kenneth
;  APPLICANT:  Blatt, Larry
;  APPLICANT:  McSwiggen, Jim
;  APPLICANT:  Morrissey, Dave
;  TITLE OF INVENTION: Method and Reagent for Inhibiting Hepatitis B Virus Replication
;  FILE REFERENCE: MBHB00-845-H (400/029)
;  CURRENT APPLICATION NUMBER: US/09/877,478
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: US 07/882,712
;  PRIOR FILING DATE: 1992-05-14
;  PRIOR APPLICATION NUMBER: US 09/531,025
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: US 09/636,385
;  PRIOR FILING DATE: 2000-08-09
;  PRIOR APPLICATION NUMBER: US 09/696,347
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 08/193,627
;  PRIOR FILING DATE: 1994-02-07
;  PRIOR APPLICATION NUMBER: US 08/433,993
;  PRIOR FILING DATE: 1995-05-04
;  PRIOR APPLICATION NUMBER: US 08/434,504
;  PRIOR FILING DATE: 1995-05-04
;  PRIOR APPLICATION NUMBER: US 09/436,430
;  PRIOR FILING DATE: 1999-11-08
;  NUMBER OF SEQ ID NOS: 6586
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4121
;   LENGTH: 37
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Enzymatic Nucleic Acid
US-09-877-478-4121

  Query Match             0.8%;  Score 18.8;  DB 9;  Length 37;
  Best Local Similarity   60.0%;  Pred. No. 1.5e+05;
  Matches   18;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy    177 AAATGAAAGCATGGACCATGGGGGAGTTGG 206
          |||:||  |||:| || |:| | || ::||
Db      6 AAAUGAUGGCAUGCACUAUGCGCGACUUGG 35


RESULT 80
US-09-877-478-3202
; Sequence 3202, Application US/09877478
; Publication No. US20030068301A1
; GENERAL INFORMATION:
;  APPLICANT: Ribozyme Pharmaceuticals, Inc.
;  APPLICANT:  Draper, Kenneth
;  APPLICANT:  Blatt, Larry
;  APPLICANT:  McSwiggen, Jim
;  APPLICANT:  Morrissey, Dave
;  TITLE OF INVENTION: Method and Reagent for Inhibiting Hepatitis B Virus Replication
;  FILE REFERENCE: MBHB00-845-H (400/029)
;  CURRENT APPLICATION NUMBER: US/09/877,478
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: US 07/882,712
;  PRIOR FILING DATE: 1992-05-14
;  PRIOR APPLICATION NUMBER: US 09/531,025
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: US 09/636,385
;  PRIOR FILING DATE: 2000-08-09
;  PRIOR APPLICATION NUMBER: US 09/696,347
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 08/193,627
;  PRIOR FILING DATE: 1994-02-07
;  PRIOR APPLICATION NUMBER: US 08/433,993
;  PRIOR FILING DATE: 1995-05-04
;  PRIOR APPLICATION NUMBER: US 08/434,504
;  PRIOR FILING DATE: 1995-05-04
;  PRIOR APPLICATION NUMBER: US 09/436,430
;  PRIOR FILING DATE: 1999-11-08
;  NUMBER OF SEQ ID NOS: 6586
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3202
;   LENGTH: 38
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Enzymatic Nucleic Acid
US-09-877-478-3202

  Query Match             0.8%;  Score 18.8;  DB 9;  Length 38;
  Best Local Similarity   55.3%;  Pred. No. 1.5e+05;
  Matches   21;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    557 GAGGAGAACTATTTATGCAGTTAGAAAGAGAGGGAATA 594
          | ||||| |:  : | || |::||   || ||||| :|
Db      1 GUGGAGAUCUGAUGAGGCCGUUAGGCCGAAAGGGAGUA 38


RESULT 81
US-09-247-890-14
; Sequence 14, Application US/09247890
; Publication No. US20020198162A1
; GENERAL INFORMATION:
;  APPLICANT: Punnonen, Juha
;  APPLICANT: Bass, Steven H.
;  APPLICANT: Whalen, Robert Gerald
;  APPLICANT: Howard, Russell
;  APPLICANT: Stemmer, Willem P.C.
;  APPLICANT: Maxygen, Inc.
;  TITLE OF INVENTION: Antigen Library Immunization
;  FILE REFERENCE: 018097-028710US
;  CURRENT APPLICATION NUMBER: US/09/247,890
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: US 60/074,294
;  EARLIER FILING DATE: 1998-02-11
;  EARLIER APPLICATION NUMBER: US 60/105,509
;  EARLIER FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 39
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:reverse HBV ayw
;   OTHER INFORMATION: PCR primer
US-09-247-890-14

  Query Match             0.8%;  Score 18.8;  DB 9;  Length 39;
  Best Local Similarity   76.7%;  Pred. No. 1.6e+05;
  Matches   23;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;
```

```
Qy   2050 GGGAAGGCAAAGACAAAAGAAACTTACCAA 2079
          ||| |  |||||||||||||| ||   ||
Db      4 GGGTACCCAAAGACAAAAGAAAATTGGTAA 33
```

RESULT 82
US-10-043-573-131/c
; Sequence 131, Application US/10043573
; Publication No. US20030032025A1
; GENERAL INFORMATION:
; APPLICANT: Lemieux, Bertrand
; Landry, Benoit S.
; Sapolsky, Ronald J.
; TITLE OF INVENTION: Brassica Polymorphisms
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/043,573
; FILING DATE: 09-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/813,507
; FILING DATE: 07-MAR-1997
; APPLICATION NUMBER: US 60/032,069
; FILING DATE: 02-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 018547-030100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415 576-0200
; TELEFAX: 415 576-0200
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 131:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 131:
US-10-043-573-131

Query Match 0.8%; Score 18.8; DB 9; Length 42;
Best Local Similarity 61.9%; Pred. No. 1.6e+05;
Matches 26; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

```
Qy    757 ATTCATGATGGAACAGTCACACACACATTTTGTGGAACAATA 798
          ||  |||     ||||||||::|||||||  |   ||   |||
Db     42 ATGGATGCAACCACAGTCACWMACACATTGTCCTGATACATA 1
```

RESULT 83
US-10-046-935-34/c
; Sequence 34, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-046-935-34

Query Match 0.8%; Score 18.8; DB 9; Length 51;
Best Local Similarity 90.9%; Pred. No. 1.9e+05;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy    430 AAAGATACAGCTCATACAAAAG 451
          ||||  ||||||||||||||||
Db     22 AAAGCCACAGCTCATACAAAAG 1
```

RESULT 84
US-09-878-178-34/c
; Sequence 34, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527
; CURRENT APPLICATION NUMBER: US/09/878,178
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 2237
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-878-178-34

Query Match 0.8%; Score 18.8; DB 9; Length 51;
Best Local Similarity 90.9%; Pred. No. 1.9e+05;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy    430 AAAGATACAGCTCATACAAAAG 451
          ||||  ||||||||||||||||
Db     22 AAAGCCACAGCTCATACAAAAG 1
```

RESULT 85
US-10-146-502-34/c
; Sequence 34, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C2
; CURRENT APPLICATION NUMBER: US/10/146,502
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-146-502-34

```
  Query Match             0.8%;  Score 18.8;  DB 9;  Length 51;
  Best Local Similarity   90.9%;  Pred. No. 1.9e+05;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    430 AAAGATACAGCTCATACAAAAG 451
          ||||  ||||||||||||||||
Db     22 AAAGCCACAGCTCATACAAAAG 1


RESULT 86
US-10-027-805-1
; Sequence 1, Application US/10027805
; Patent No. US20020164725A1
;   GENERAL INFORMATION:
;        APPLICANT: ROBERTSON, Daniel E.
;                   MURPHY, Dennis
;                   REID, John
;                   MAFFIA, Anthony
;                   LINK, Steven
;                   SWANSON, Ronald V.
;                   WARREN, Patrick V.
;                   KOSMOTKA, Anna
;        TITLE OF INVENTION: ESTERASES
;        NUMBER OF SEQUENCES: 42
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: FISH & RICHARDSON P.C.
;             STREET: 4225 EXECUTIVE SQUARE, STE 1400
;             CITY: LA JOLLA
;             STATE: CALIFORNIA
;             COUNTRY: USA
;             ZIP: 92037
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5 INCH DISKETTE
;             COMPUTER: IBM PS/2
;             OPERATING SYSTEM: MS-DOS
;             SOFTWARE: WORD PERFECT 6.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/027,805
;             FILING DATE: 21-Dec-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/602,359
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: HAILE, LISA A.
;             REGISTRATION NUMBER: 38,347
;             REFERENCE/DOCKET NUMBER: 09010/010001
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 619-678-5070
;             TELEFAX: 619-678-5099
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 52 NUCLEOTIDES
;             TYPE: NUCLEIC ACID
;             STRANDEDNESS: SINGLE
;             TOPOLOGY: LINEAR
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-027-805-1

  Query Match             0.8%;  Score 18.8;  DB 9;  Length 52;
  Best Local Similarity   68.4%;  Pred. No. 1.9e+05;
  Matches   26;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   1734 ACACAGGAAAAATAAACGTGGATTTTAAAAAATCAATC 1771
          | | |||| |||| |||   |  |||||| || || ||
Db     14 AAAGAGGAGAAATTAACTATGTCTTTAAACAAGCACTC 51


RESULT 87
US-09-907-900-58/c
; Sequence 58, Application US/09907900
; Patent No. US20020172997A1
; GENERAL INFORMATION:
;  APPLICANT: Hartley, James L.
;  APPLICANT:  Brasch, Michael A.
;  APPLICANT:  Temple, Gary F.
;  APPLICANT:  Fox, Donna K.
;  TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having
;  TITLE OF INVENTION:  Recombination Sites
;  FILE REFERENCE: 0942.2850004
;  CURRENT APPLICATION NUMBER: US/09/907,900
;  CURRENT FILING DATE:  2001-07-19
;  PRIOR APPLICATION NUMBER: 09/177,387
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
;   LENGTH: 52
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: oligonucleotide
US-09-907-900-58

  Query Match             0.8%;  Score 18.8;  DB 9;  Length 52;
  Best Local Similarity   63.0%;  Pred. No. 1.9e+05;
  Matches   29;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   1995 TGCAACTAAGGACAAATTAGCATGCAAGCTTGGTCAAACTTTTCCC 2040
          | ||  |  | |   |||||| |||    |||  ||||||| ||||
Db     47 TACATATTTGAATGTATTAGCCTGCTTTTTTGTACAAACTTGTCCC 2


RESULT 88
US-09-907-719-58/c
; Sequence 58, Application US/09907719
; Publication No. US20020192819A1
; GENERAL INFORMATION:
;  APPLICANT: Hartley, James L.
;  APPLICANT:  Brasch, Michael A.
;  APPLICANT:  Temple, Gary F.
;  APPLICANT:  Fox, Donna K.
;  TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having
;  TITLE OF INVENTION:  Recombination Sites
;  FILE REFERENCE: 0942.2850004
;  CURRENT APPLICATION NUMBER: US/09/907,719
;  CURRENT FILING DATE:  2001-07-19
;  PRIOR APPLICATION NUMBER: US/09/177,387
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
;   LENGTH: 52
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: oligonucleotide
US-09-907-719-58

  Query Match             0.8%;  Score 18.8;  DB 9;  Length 52;
  Best Local Similarity   63.0%;  Pred. No. 1.9e+05;
  Matches   29;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   1995 TGCAACTAAGGACAAATTAGCATGCAAGCTTGGTCAAACTTTTCCC 2040
          | ||  |  | |   |||||| |||    |||  ||||||| ||||
Db     47 TACATATTTGAATGTATTAGCCTGCTTTTTTGTACAAACTTGTCCC 2


RESULT 89
US-10-027-804-1
; Sequence 1, Application US/10027804
; Publication No. US20030054530A1
```

```
;   GENERAL INFORMATION:
;        APPLICANT: ROBERTSON, Daniel E.
;                   MURPHY, Dennis
;                   REID, John
;                   MAFFIA, Anthony
;                   LINK, Steven
;                   SWANSON, Ronald V.
;                   WARREN, Patrick V.
;                   KOSMOTKA, Anna
;        TITLE OF INVENTION: ESTERASES
;        NUMBER OF SEQUENCES: 42
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: FISH & RICHARDSON P.C.
;             STREET: 4225 EXECUTIVE SQUARE, STE 1400
;             CITY: LA JOLLA
;             STATE: CALIFORNIA
;             COUNTRY: USA
;             ZIP: 92037
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5 INCH DISKETTE
;             COMPUTER: IBM PS/2
;             OPERATING SYSTEM: MS-DOS
;             SOFTWARE: WORD PERFECT 6.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/027,804
;             FILING DATE: 21-Dec-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/602,359
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: HAILE, LISA A.
;             REGISTRATION NUMBER: 38,347
;             REFERENCE/DOCKET NUMBER: 09010/010001
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 619-678-5070
;             TELEFAX: 619-678-5099
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 52 NUCLEOTIDES
;             TYPE: NUCLEIC ACID
;             STRANDEDNESS: SINGLE
;             TOPOLOGY: LINEAR
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-027-804-1

  Query Match             0.8%;  Score 18.8;  DB 9;  Length 52;
  Best Local Similarity   68.4%;  Pred. No. 1.9e+05;
  Matches   26;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   1734 ACACAGGAAAAATAAACGTGGATTTTAAAAAATCAATC 1771
          | | |||| |||| |||   |  |||||| || || ||
Db     14 AAAGAGGAGAAATTAACTATGTCTTTAAACAAGCACTC 51


RESULT 90
US-09-855-797A-58/c
; Sequence 58, Application US/09855797A
; Patent No. US20020094574A1
; GENERAL INFORMATION:
;  APPLICANT: Hartley, James L.
;  APPLICANT:  Brasch, Michael A.
;  APPLICANT:  Temple, Gary F.
;  APPLICANT:  Fox, Donna K.
;  TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having
;  TITLE OF INVENTION:  Recombination Sites
;  FILE REFERENCE: 0942.2850008
;  CURRENT APPLICATION NUMBER: US/09/855,797A
;  CURRENT FILING DATE:  2001-05-16
;  PRIOR APPLICATION NUMBER: 09/296,281
;  PRIOR FILING DATE: 1999-04-22
;  PRIOR APPLICATION NUMBER: US 60/065,930
;  PRIOR FILING DATE: 1997-10-24
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
;    LENGTH: 52
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence: synthetic
;    OTHER INFORMATION: oligonucleotide
US-09-855-797A-58

  Query Match             0.8%;  Score 18.8;  DB 10;  Length 52;
  Best Local Similarity   63.0%;  Pred. No. 1.9e+05;
  Matches   29;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   1995 TGCAACTAAGGACAAATTAGCATGCAAGCTTGGTCAAACTTTTCCC 2040
          | ||  |  | |   |||||| |||    |||  ||||||| ||||
Db     47 TACATATTTGAATGTATTAGCCTGCTTTTTTGTACAAACTTGTCCC 2


RESULT 91
US-09-903-410-1
; Sequence 1, Application US/09903410
; Patent No. US20020146799A1
; GENERAL INFORMATION:
;  APPLICANT: DIVERSA CORPORATION
;  APPLICANT:  ROBERTSON, Dan
;  APPLICANT:  MURPHY, Dennis
;  APPLICANT:  REID, John
;  APPLICANT:  MAFFIA, Anthony
;  APPLICANT:  LINK, Steven
;  APPLICANT:  SWANSON, Ronald
;  APPLICANT:  WARREN, Patrick
;  APPLICANT:  KOSMOTKA, Anna
;  TITLE OF INVENTION: ENZYMES HAVING ESTERASE ACTIVITY AND METHODS OF USE THEREOF
;  FILE REFERENCE: DIVER1180-2
;  CURRENT APPLICATION NUMBER: US/09/903,410
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: US 09/382,242
;  PRIOR FILING DATE: 1999-08-24
;  PRIOR APPLICATION NUMBER: US 08/602,359
;  PRIOR FILING DATE: 1996-02-16
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;    LENGTH: 52
;    TYPE: DNA
;    ORGANISM: Artificial sequence
;    FEATURE:
;    OTHER INFORMATION: Primer for PCR
US-09-903-410-1

  Query Match             0.8%;  Score 18.8;  DB 10;  Length 52;
  Best Local Similarity   68.4%;  Pred. No. 1.9e+05;
  Matches   26;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   1734 ACACAGGAAAAATAAACGTGGATTTTAAAAAATCAATC 1771
          | | |||| |||| |||   |  |||||| || || ||
Db     14 AAAGAGGAGAAATTAACTATGTCTTTAAACAAGCACTC 51


RESULT 92
US-09-995-225-51
; Sequence 51, Application US/09995225
; Publication No. US20020193584A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, Ruoping
;  APPLICANT:  Chu, Zhi Liang
;  APPLICANT:  Dang, Huong T.
;  APPLICANT:  Lowitz, Kevin P.
;  APPLICANT:  Pride, Cameron
;  TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human G
```

```
;  TITLE OF INVENTION:  Receptors
;  FILE REFERENCE: AREN-0308
;  CURRENT APPLICATION NUMBER: US/09/995,225
;  CURRENT FILING DATE:  2001-11-26
;  PRIOR APPLICATION NUMBER: 09/170,496
;  PRIOR FILING DATE: 1998-10-13
;  PRIOR APPLICATION NUMBER: PCT/US99/23938
;  PRIOR FILING DATE: 1998-10-13
;  PRIOR APPLICATION NUMBER: 60/253,404
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/255,366
;  PRIOR FILING DATE: 2000-12-12
;  PRIOR APPLICATION NUMBER: 60/270,286
;  PRIOR FILING DATE: 2001-02-20
;  PRIOR APPLICATION NUMBER: 60/282,365
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: 60/270,266
;  PRIOR FILING DATE: 2001-02-20
;  PRIOR APPLICATION NUMBER: 60/282,032
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: 60/282,358
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: 60/282,356
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: 60/290,917
;  PRIOR FILING DATE: 2001-05-14
;  PRIOR APPLICATION NUMBER: 60/309,208
;  PRIOR FILING DATE: 2001-07-31
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51
;   LENGTH: 53
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: No. US20020193584A1el Sequence
US-09-995-225-51

  Query Match             0.8%;  Score 18.8;  DB 9;  Length 53;
  Best Local Similarity   68.4%;  Pred. No. 1.9e+05;
  Matches   26;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    132 AGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGG 169
          | ||| ||| | |  | ||| | ||||| |||| || |
Db     16 ACAGGCCGCAGTCCTTCAGGTTCAGCTGCAGGATGGTG 53


RESULT 93
US-10-083-168-92
; Sequence 92, Application US/10083168
; Publication No. US20030023069A1
; GENERAL INFORMATION:
;  APPLICANT: Liaw, Chen W.
;  APPLICANT:  Chalmers, Derek T.
;  APPLICANT:  Behan, Dominic P.
;  APPLICANT:  Maciejewski-Lenior, Dominique
;  APPLICANT:  Leonard, James N.
;  APPLICANT:  Ortuno, Daniel
;  APPLICANT:  Lin, I-Lin
;  TITLE OF INVENTION: Endogenous And No. US20030023069A1-Endogenous, Constitutively Act
;  TITLE OF INVENTION:  Receptors
;  FILE REFERENCE: AREN-0320
;  CURRENT APPLICATION NUMBER: US/10/083,168
;  CURRENT FILING DATE:  2002-02-26
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 92
;   LENGTH: 53
;   TYPE: DNA
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: No. US20030023069A1el Sequence
US-10-083-168-92

  Query Match             0.8%;  Score 18.8;  DB 9;  Length 53;
  Best Local Similarity   68.4%;  Pred. No. 1.9e+05;
  Matches   26;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    132 AGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGG 169
          | ||| ||| | |  | ||| | ||||| |||| || |
Db     16 ACAGGCCGCAGTCCTTCAGGTTCAGCTGCAGGATGGTG 53


RESULT 94
US-08-781-986A-2724/c
; Sequence 2724, Application US/08781986A
; Publication No. US20030054436A1
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  2724:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  54 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-781-986A-2724

  Query Match             0.8%;  Score 18.8;  DB 7;  Length 54;
  Best Local Similarity   63.0%;  Pred. No. 1.9e+05;
  Matches   29;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   1016 GAAATGCTGCTTCTCGTCTGGGAGCTGGTCCTGGGGACGCTGGAGA 1061
          ||| | ||||| | ||    | || ||  || ||  | ||||| ||
Db     54 GAATTCCTGCTGCACGCTCTGCAGTTGTCCCGGGCCAAGCTGGTGA 9


RESULT 95
US-09-944-036-30
; Sequence 30, Application US/09944036
; Patent No. US20020055095A1
; GENERAL INFORMATION:
;  APPLICANT: YANG, Yeasing Y.
;  APPLICANT:  BRENTANO, Steven T.
;  APPLICANT:  BABOLA, Odile
;  APPLICANT:  TRAN, Nathalie
;  APPLICANT:  VERNET, Guy
;  TITLE OF INVENTION: AMPLIFICATION OF HIV-1 SEQUENCES FOR DETECTION OF
;  TITLE OF INVENTION:  SEQUENCES ASSOCIATED WITH DRUG-RESISTANCE MUTATIONS
```

```
;  FILE REFERENCE: GP114-02.UT
;  CURRENT APPLICATION NUMBER: US/09/944,036
;  CURRENT FILING DATE:  2001-08-31
;  PRIOR APPLICATION NUMBER: US 60/229,790
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
;   LENGTH: 56
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Capture
;   OTHER INFORMATION: oligonucleotide
US-09-944-036-30

  Query Match             0.8%;  Score 18.8;  DB 10;  Length 56;
  Best Local Similarity   63.0%;  Pred. No. 2e+05;
  Matches   29;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   2220 ACAACTTCTGTTTCTTCTCTTGGTGAAATAATAAAATGCAAATGAA 2265
           | ||||||||| |||| | |     ||| || ||||   |||. ||
Db      7 ATAACTTCTGCTTCTATTTTAAAAAAAAAAAAAAAAAAAAAAAAA 52


RESULT 96
US-08-781-986A-1941
; Sequence 1941, Application US/08781986A
; Publication No. US20030054436A1
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  1941:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  57 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-781-986A-1941

  Query Match             0.8%;  Score 18.8;  DB 7;  Length 57;
  Best Local Similarity   63.0%;  Pred. No. 2e+05;
  Matches   29;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy    374 CTGGGAAAATATTTGCCATGAAGGTGCTTAAAAAGGCAATGATAGT 419
          | |||||||| ||||  || |||  |   |||||  |  | | | |
Db      1 CCGGGAAAATTTTTGGAATTAAGTGGAAAAAAAATCCCCTTAAAAT 46


RESULT 97
US-09-981-803-29/c
; Sequence 29, Application US/09981803
; Publication No. US20030032092A1
; GENERAL INFORMATION:
;  APPLICANT: Joel CROUZET
;  APPLICANT:  Daniel SCHERMAN
;  APPLICANT:  Beatrice CAMERON
;  APPLICANT:  Pierre WILS
;  APPLICANT:  Anne-Marie DARQUET
;  TITLE OF INVENTION: DNA MOLECULES, PREPARATION AND USE IN GENE THERAPY
;  FILE REFERENCE: MINICIRCLE
;  CURRENT APPLICATION NUMBER: US/09/981,803
;  CURRENT FILING DATE:  2001-10-19
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
;   LENGTH: 57
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Description of the artificial sequence:
;   OTHER INFORMATION: oligonucleotide
US-09-981-803-29

  Query Match             0.8%;  Score 18.8;  DB 9;  Length 57;
  Best Local Similarity   59.3%;  Pred. No. 2e+05;
  Matches   32;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy   1487 CAATGAGTGGGGAAGCATCGGCACCACTTCCAATACGACAGCCGAACTCTGGGC 1540
           | |  |||||  | ||||| || | | ||   |||  | ||| |   || | ||
Db     55 CTACTAGTGGCCATGCATCCGCTCAAGTTAGTATAAAAAAGCAGGCTTCAGAGC 2


RESULT 98
US-09-878-574-9023/c
; Sequence 9023, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 9023
;   LENGTH: 58
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: 701101993H1
US-09-878-574-9023

  Query Match             0.8%;  Score 18.8;  DB 10;  Length 58;
  Best Local Similarity   59.3%;  Pred. No. 2e+05;
  Matches   32;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy    192 CCATGGGGGAGTTGGACCATATGAACTTGGCATGGAACATTGTGAGAAATTTGA 245
            ||   ||| |    ||| || |  ||  ||| | | | | ||||||| || | |
Db     57 CCTCCGGGCACGGCGACGATTTTGACAAGGCTTTGGAAAGTGTGAGACATGTCA 4


RESULT 99
US-10-007-968-8/c
; Sequence 8, Application US/10007968
; Patent No. US20020159977A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Couto, Linda B.
;  APPLICANT:  Colosi, Peter C.
;  TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII
;  TITLE OF INVENTION:  by Target Cells
;  FILE REFERENCE: Avigen-04082
;  CURRENT APPLICATION NUMBER: US/10/007,968
;  CURRENT FILING DATE:  2001-12-13
;  PRIOR APPLICATION NUMBER: 09/740,211
;  PRIOR FILING DATE: 2000-12-18
;  PRIOR APPLICATION NUMBER: 60/125,974
;  PRIOR FILING DATE: 1999-03-24
;  PRIOR APPLICATION NUMBER: 60/104,994
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 59
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
US-10-007-968-8

  Query Match             0.8%;  Score 18.8;  DB 9;  Length 59;
  Best Local Similarity   63.0%;  Pred. No. 2.1e+05;
  Matches   29;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   2237 TCTTGGTGAAATAATAAAATGCAAATGAATCATTGTTAACACAGCT 2282
          | ||||| ||| ||  | ||   |  || || || |  |||||| |
Db     55 TATTGGTTAAAGAAGTATATTAGAGCGAGTCTTTCTGCACACAGAT 10


RESULT 100
US-10-007-280A-2
; Sequence 2, Application US/10007280A
; Publication No. US20030059784A1
; GENERAL INFORMATION:
;  APPLICANT: Sun, Yongming
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Salceda, Susana
;  APPLICANT:  Chenghua, Liu
;  TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro
;  FILE REFERENCE: DEX-0257
;  CURRENT APPLICATION NUMBER: US/10/007,280A
;  CURRENT FILING DATE:  2001-11-07
;  PRIOR APPLICATION NUMBER: US 60/246,640
;  PRIOR FILING DATE: 2000-11-08
;  NUMBER OF SEQ ID NOS: 238
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 59
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-007-280A-2

  Query Match             0.8%;  Score 18.8;  DB 9;  Length 59;
  Best Local Similarity   63.0%;  Pred. No. 2.1e+05;
  Matches   29;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   1645 AGAGGGAGATGTGTGAGCATCCTGCAAGGTGAAACAAGACTCAAAA 1690
          | | | |||| ||  ||||  || ||   | |||||   |||| ||
Db      5 AAATGTAGATATGAAAGCAGACTACAGTATAAAACACTGCTCAGAA 50


Search completed: April 19, 2003, 12:08:13
Job time : 248 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: April 19, 2003, 06:54:23 ; Search time 3716 Seconds
(without alignments)
10224.600 Million cell updates/sec

Title: US-09-920-677-3
Perfect score: 2346
Sequence: 1 gcacgaggctgcggcgggtc.....gtaacccagctgcggagcct 2346

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 146654

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	37.4	1.6	47	9	AA153855	AA153855 mq46f01.r
c	2	26.6	1.1	48	13	BJ039490	BJ039490 BJ039490
	3	25.2	1.1	58	17	B02929	B02929 cSRL-163B1-
	4	24.4	1.0	54	9	AU269582	AU269582 AU269582
c	5	23.8	1.0	45	9	AU266762	AU266762 AU266762
c	6	23.4	1.0	42	17	TA261E04Q	AL488612 T. brucei
	7	23.4	1.0	56	9	AA761128	AA761128 nz11e03.s
	8	23.4	1.0	58	9	AA107349	AA107349 ml93a10.r
c	9	23.2	1.0	58	14	BQ907332	BQ907332 P004C10 O
	10	23.2	1.0	59	17	BH813384	BH813384 SALK_0640
	11	23	1.0	50	12	BG271427	BG271427 nai50f05.
	12	23	1.0	51	17	AZ508445	AZ508445 1M0350F11
	13	23	1.0	59	17	AZ621827	AZ621827 1M0455C10
c	14	23	1.0	60	9	AA554929	AA554929 nk31g05.s
	15	23	1.0	60	10	AW411378	AW411378 fh12a03.y
	16	23	1.0	60	12	BG409218	BG409218 gb88b10.y
c	17	22.8	1.0	52	10	AW249993	AW249993 2819345.3
c	18	22.8	1.0	53	14	BM880091	BM880091 ku04h02.y
c	19	22.8	1.0	54	17	AZ482068	AZ482068 1M0306M23
c	20	22.8	1.0	59	9	AU264598	AU264598 AU264598
c	21	22.8	1.0	59	9	AU267788	AU267788 AU267788
c	22	22.8	1.0	59	14	BQ397088	BQ397088 NISC_ng25
c	23	22.6	1.0	51	9	AU266759	AU266759 AU266759
c	24	22.6	1.0	53	9	AU006675	AU006675 AU006675
	25	22.6	1.0	56	9	AU264938	AU264938 AU264938
	26	22.6	1.0	56	17	B03428	B03428 cSRL-179A8-
c	27	22.6	1.0	57	9	AA815875	AA815875 vp74b01.r
	28	22.6	1.0	58	17	B00931	B00931 cSRL-122a1-
	29	22.6	1.0	60	17	CNS06CC9	AL392335 T3 end of
c	30	22.4	1.0	59	9	AU265345	AU265345 AU265345
	31	22.4	1.0	60	9	AU263817	AU263817 AU263817
c	32	22.2	0.9	44	9	AU268768	AU268768 AU268768
c	33	22.2	0.9	44	9	AU268819	AU268819 AU268819
	34	22.2	0.9	52	12	BF643229	BF643229 NF002B08E
c	35	22.2	0.9	53	17	TA385D03Q	AL498857 T. brucei
c	36	22.2	0.9	54	9	AA937418	AA937418 oj09a06.s
	37	22.2	0.9	55	17	AZ760065	AZ760065 1M0553B02
	38	22.2	0.9	56	13	BI088641	BI088641 602851403
	39	22.2	0.9	57	9	AU270417	AU270417 AU270417
c	40	22.2	0.9	58	13	BG943501	BG943501 ax38e11.x
c	41	22.2	0.9	59	13	BG943837	BG943837 ax42f06.x
	42	22	0.9	49	9	AA014672	AA014672 mh30b01.r
c	43	22	0.9	51	12	BG301974	BG301974 kt07h05.y
c	44	22	0.9	55	9	AI168523	AI168523 ox10h11.s
	45	22	0.9	57	9	AU264614	AU264614 AU264614
c	46	22	0.9	57	13	BI745330	BI745330 rk99b01.y
c	47	22	0.9	58	13	BI842365	BI842365 fq82g06.y
	48	22	0.9	58	14	BM965520	BM965520 ko13a10.y
	49	22	0.9	58	17	AL762620	AL762620 Arabidops
	50	22	0.9	60	9	AL779773	AL779773 AL779773
	51	21.8	0.9	38	13	BJ059311	BJ059311 BJ059311
	52	21.8	0.9	46	13	BJ057533	BJ057533 BJ057533
c	53	21.8	0.9	48	9	AU263470	AU263470 AU263470
	54	21.8	0.9	48	10	BE545959	BE545959 601069318
	55	21.8	0.9	49	9	AA920052	AA920052 vy63h11.r
	56	21.8	0.9	52	9	AA823664	AA823664 vr69d09.s
	57	21.8	0.9	52	17	BH792320	BH792320 SALK_0634
	58	21.8	0.9	54	9	AU258750	AU258750 AU258750
c	59	21.8	0.9	55	9	AI888083	AI888083 wm29g07.x
	60	21.8	0.9	55	17	AZ331030	AZ331030 1M0056G09
	61	21.8	0.9	58	12	BF791590	BF791590 602251706
c	62	21.8	0.9	58	13	BI846606	BI846606 fq80g06.x
c	63	21.8	0.9	59	9	AA726911	AA726911 vu37b05.r
	64	21.8	0.9	60	9	AA547914	AA547914 MB3D6V2G0
	65	21.6	0.9	45	12	BE896253	BE896253 601438976
	66	21.6	0.9	50	9	AU107554	AU107554 AU107554
c	67	21.6	0.9	53	9	AU009949	AU009949 AU009949
	68	21.6	0.9	54	14	BQ565734	BQ565734 gi44e07.y
	69	21.6	0.9	54	17	BH846951	BH846951 SALK_0121
	70	21.6	0.9	55	9	AL786691	AL786691 AL786691
c	71	21.6	0.9	56	9	AU264938	AU264938 AU264938
c	72	21.6	0.9	56	13	BI855869	BI855869 603383029
c	73	21.6	0.9	58	13	BI705223	BI705223 fr66e10.y
c	74	21.6	0.9	58	13	BI709260	BI709260 fs62b03.y
	75	21.6	0.9	59	13	BI496827	BI496827 df128f02.
c	76	21.6	0.9	60	12	BG348123	BG348123 de78g05.y
c	77	21.6	0.9	60	14	C25778	C25778 C25778 Dict
c	78	21.6	0.9	60	14	N81844	N81844 TgESTzy48g0
c	79	21.4	0.9	41	17	AZ368409	AZ368409 1M0118A04

	80	21.4	0.9	46	9	AI431480 AI431480 th40f03.x
	81	21.4	0.9	49	12	BG222966 BG222966 nah40b02.
c	82	21.4	0.9	51	13	BJ064587 BJ064587 BJ064587
	83	21.4	0.9	52	12	BG271387 BG271387 nai49h05.
	84	21.4	0.9	53	12	BG231005 BG231005 nai43b05.
	85	21.4	0.9	55	17	AZ366215 AZ366215 1M0115A14
c	86	21.4	0.9	57	14	N63590 N63590 yz77a05.s1
c	87	21.4	0.9	57	17	B03506 B03506 cSRL-180D5-
c	88	21.4	0.9	59	9	AI473426 AI473426 tj13g01.x
c	89	21.4	0.9	59	17	AL769947 AL769947 Arabidops
c	90	21.2	0.9	39	10	AV837933 AV837933 AV837933
	91	21.2	0.9	39	17	BH235137 BH235137 MSAD_E07.
c	92	21.2	0.9	46	9	AA027616 AA027616 mi08c08.r
	93	21.2	0.9	49	12	BG271415 BG271415 nai50d11.
	94	21.2	0.9	49	12	BG819402 BG819402 602781632
	95	21.2	0.9	51	13	BJ032647 BJ032647 BJ032647
c	96	21.2	0.9	51	17	AZ948877 AZ948877 2M0212P06
c	97	21.2	0.9	52	9	AL587783 AL587783 AL587783
	98	21.2	0.9	53	17	AZ862273 AZ862273 2M0169H19
c	99	21.2	0.9	53	17	CNS013L5 AL102995 Drosophil
c	100	21.2	0.9	54	12	BG022522 BG022522 daa76h12.

ALIGNMENTS

RESULT 1
AA153855
LOCUS AA153855 47 bp mRNA linear EST 19-FEB-1997
DEFINITION mq46f01.r1 Soares_thymus_2NbMT Mus musculus cDNA clone IMAGE:581785 5' similar to gb:M60724 RIBOSOMAL PROTEIN S6 KINASE (HUMAN);, mRNA sequence.
ACCESSION AA153855
VERSION AA153855.1 GI:1725637
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 47)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:356433
Trace considered overall poor quality
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .47
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:581785"
/clone_lib="Soares_thymus_2NbMT"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 15 a 8 c 7 g 17 t
ORIGIN

Query Match 1.6%; Score 37.4; DB 9; Length 47;
Best Local Similarity 87.2%; Pred. No. 6.3e+02;
Matches 41; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1990 ATACTTGCAACTAAGGACAAATTAGCATGCAAGCTTGGTCAAACTTT 2036
||||||| |||||||||| ||||||| |||||| | ||||||||||
Db 1 ATACTTGTAACTAAGGACTAATTAGCTTGCAAGTCTTGTCAAACTTT 47

RESULT 2
BJ039490/c
LOCUS BJ039490 48 bp mRNA linear EST 05-DEC-2001
DEFINITION BJ039490 NIBB Mochii normalized Xenopus neurula library Xenopus laevis cDNA clone XL049k03 5', mRNA sequence.
ACCESSION BJ039490
VERSION BJ039490.1 GI:17372298
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 48)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES Location/Qualifiers
source 1. .48
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL049k03"
/clone_lib="NIBB Mochii normalized Xenopus neurula library"
/tissue_type="whole embryo"
/dev_stage="stage 15"
/note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Nieuwkoop and Faber. Library is subtracted and was constructed by N. Garrett and A.M. Zorn, (Wellcome/CRC Institute). "
BASE COUNT 7 a 3 c 2 g 29 t 7 others
ORIGIN

Query Match 1.1%; Score 26.6; DB 13; Length 48;
Best Local Similarity 66.7%; Pred. No. 1.8e+05;
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1741 AAAAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAA 1788
||||| | | | |||||||||| || | | | |||||||||
Db 48 AAAAAAANNCCTTTNTTTTAAAAAAAAAAANNANNGGGAAAAAAAAAA 1

RESULT 3
B02929

```
LOCUS       B02929                    58 bp    DNA     linear   GSS 13-JUL-1996
DEFINITION  cSRL-163B1-u cSRL flow sorted Chromosome 11 specific cosmid Homo
            sapiens genomic clone cSRL-163B1, DNA sequence.
ACCESSION   B02929
VERSION     B02929.1  GI:1412207
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Evans,G.A., Burbee,D., Davies,C., Hahner,L., Oliver,T., Gilbert,M.,
            Jones,D., Ward,T., Gillilan,E., Schagemann,J., Probst,S. ., Harris
            ,J., DeFord,J., McFarland,J., Burzinski,K., Khan,M., Kupfer,K. and
            Garner,H.R.
  TITLE     Genomic Sequence Sampled Map of Chromosome 11
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Evans GA, Shane Probst
            McDermott Center for Human Growth and Development
            University of Texas Southwestern Medical Center At Dallas
            5323 Harry Hines Blvd, Dallas TX 75235-8591
            Tel: 214-648-1600
            Fax: 214-648-1666
            Email: gevans@utsw.swmed.edu, shane@mcdermott.swmed.edu
            Seq primer: T7
            Class: cosmid ends
            High quality sequence stop: 58.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="cSRL-163B1"
                     /clone_lib="cSRL flow sorted Chromosome 11 specific
                     cosmid"
                     /sex="female"
                     /cell_type="chimeric hamster somatic cell hybrid"
                     /note="Vector: sCos-1; Human Chromosome 11 specific cosmid
                     library prepared from flow sorted human  Chromosome 11
                     derived from Chinese Hampster Ovary (CHO) monochromosomal
                     somatic cell hybrid, J1"
BASE COUNT       27 a      0 c      1 g     29 t      1 others
ORIGIN

  Query Match             1.1%;  Score 25.2;  DB 17;  Length 58;
  Best Local Similarity   70.2%;  Pred. No. 3.8e+05;
  Matches   33;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   2106 AATCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
           ||| | ||| |||||  | || |||||||  | | | | |||| |||
Db      1 AATTTATTTATTATAATATTAGATATTTTNAATAAAAAATTTTAATT 47


RESULT 4
AU269582
LOCUS       AU269582                  54 bp    mRNA    linear   EST 10-MAY-2002
DEFINITION  AU269582 VS Dictyostelium discoideum cDNA clone VSJ214 5', mRNA
            sequence.
ACCESSION   AU269582
VERSION     AU269582.1  GI:20528380
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Urushihara,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
            Takeuchi,I., Kohara,Y. and Tanaka,Y.
  TITLE     Population analysis of cDNAs from unicellular and multicellular
            stages of Dictyostelium discoideum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="VSJ214"
                     /clone_lib="VS"
                     /sex="mat A"
                     /dev_stage="vegetative"
BASE COUNT       40 a      3 c      2 g      9 t
ORIGIN

  Query Match             1.0%;  Score 24.4;  DB 9;  Length 54;
  Best Local Similarity   73.8%;  Pred. No. 5.8e+05;
  Matches   31;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   1759 TAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAA 1800
           ||||  |  ||  || ||| ||||||||||| |||||  ||||
Db     12 TAAACCAAAAAAAAAAGGTTCAAAAAAAAAATTAAAAAAAAA 53


RESULT 5
AU266762/c
LOCUS       AU266762                  45 bp    mRNA    linear   EST 10-MAY-2002
DEFINITION  AU266762 VS Dictyostelium discoideum cDNA clone VSG731 5', mRNA
            sequence.
ACCESSION   AU266762
VERSION     AU266762.1  GI:20525560
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Urushihara,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
            Takeuchi,I., Kohara,Y. and Tanaka,Y.
  TITLE     Population analysis of cDNAs from unicellular and multicellular
            stages of Dictyostelium discoideum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="VSG731"
                     /clone_lib="VS"
                     /sex="mat A"
                     /dev_stage="vegetative"
BASE COUNT       29 a      1 c      0 g     15 t
ORIGIN

  Query Match             1.0%;  Score 23.8;  DB 9;  Length 45;
  Best Local Similarity   72.1%;  Pred. No. 7.9e+05;
  Matches   31;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   2110 TTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
           ||||||||| |||||| | | | |||| |  ||  | |||| ||
Db     45 TTTTTTTTTTATAAAAAAAAAATTTTAATTTATTTATTTGTTT 3


RESULT 6
TA261E04Q/c
LOCUS       TA261E04Q                 42 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 261e04, reverse sequence,
```

```
            genomic survey sequence.
ACCESSION   AL488612
VERSION     AL488612.1  GI:11863920
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="261e04"
BASE COUNT       15 a      3 c      4 g     20 t
ORIGIN

  Query Match             1.0%;  Score 23.4;  DB 17;  Length 42;
  Best Local Similarity   73.2%;  Pred. No. 9.8e+05;
  Matches   30;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   2096 AAACAACCTGAATCTTTTTTTTATATAAATATATATTTTTC 2136
          ||| || | ||||| | | | |||||| || |||||  |||
Db     42 AAAAAAACAGAATCATCTATATATATATATTTATATAGTTC 2


RESULT 7
AA761128
LOCUS       AA761128                 56 bp    mRNA    linear   EST 07-FEB-1998
DEFINITION  nz11e03.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1287484 3',
            mRNA sequence.
ACCESSION   AA761128
VERSION     AA761128.1  GI:2810058
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1195   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 51.
FEATURES             Location/Qualifiers
     source          1. .56
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1287484"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTT-3'
                     ]. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT       18 a      2 c      7 g     29 t
ORIGIN

  Query Match             1.0%;  Score 23.4;  DB 9;  Length 56;
  Best Local Similarity   73.2%;  Pred. No. 9.7e+05;
  Matches   30;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTT 2148
           |||||||||   ||||| || | |||  || | ||||||
Db      4 TTTTTTTTTAGTAAAATAAATGTATTTTTAAAATATTTTT 44


RESULT 8
AA107349
LOCUS       AA107349                 58 bp    mRNA    linear   EST 04-FEB-1997
DEFINITION  ml93a10.r1 Stratagene mouse kidney (#937315) Mus musculus cDNA
            clone IMAGE:519546 5' similar to SW:ATP6_MOUSE P00848 ATP SYNTHASE
            A CHAIN ;, mRNA sequence.
ACCESSION   AA107349
VERSION     AA107349.1  GI:1658632
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:313394
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
```

```
     source          1. .58
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:519546"
                     /clone_lib="Stratagene mouse kidney (#937315)"
                     /sex="females"
                     /tissue_type="kidney"
                     /dev_stage="4 weeks"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: kidney; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Average insert size: 1.0 kb; Uni-ZAP XR Vector;
                     ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       17 a      14 c       8 g      19 t
ORIGIN

  Query Match             1.0%;  Score 23.4;  DB 9;  Length 58;
  Best Local Similarity   67.3%;  Pred. No. 9.7e+05;
  Matches   33;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    1793 AAGCAAAATAGTATTGCTGAACTCTTAGGCACATCAATTAATTGATTCC 1841
           || |||||| |||||    | || |   | || |||||| | ||||||
Db       2 AACCAAAATCCTATTGGCTTATTCATTACCCCAACAATTATTGGATTCC 50


RESULT 9
BQ907332/c
LOCUS       BQ907332                  58 bp    mRNA    linear   EST 19-AUG-2002
DEFINITION  P004C10 Oryza sativa mature leaf library induced by M.grisea Oryza
            sativa cDNA clone P004C10, mRNA sequence.
ACCESSION   BQ907332
VERSION     BQ907332.1  GI:22306110
KEYWORDS    EST.
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X., Wu
            ,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.
  TITLE     A Gene Expression Screen in Oryza sativa
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Dong HT
            Laboratory of Functional Genetics
            Bio-technology Institute of Zhejiang University
            Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
            Tel: 0086-571-86892051
            Fax: 0086-571-86961525
            Email: htdong@zjuem.zju.edu.cn
            Seq primer: M13 forward primer.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Oryza sativa"
                     /db_xref="taxon:4530"
                     /clone="P004C10"
                     /clone_lib="Oryza sativa mature leaf library induced by
                     M.grisea"
                     /tissue_type="leaf"
                     /dev_stage="Mature stage"
                     /note="Vector: pSport2"
BASE COUNT       33 a      4 c       1 g      19 t      1 others
ORIGIN

  Query Match             1.0%;  Score 23.2;  DB 14;  Length 58;
  Best Local Similarity   64.2%;  Pred. No. 1.1e+06;
  Matches   34;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy    2225 TTCTGTTTCTTCTCTTGGTGAAATAATAAAATGCAAATGAATCATTGTTAACA 2277
           || | ||| || | ||    || ||||  | || | || ||| |||| ||| |
Db      55 TTTTTTTTTTTTTTTTAAATAATTAATTTACTGNATATAAATTATTGATAAAA 3


RESULT 10
BH813384
LOCUS       BH813384                  59 bp    DNA     linear   GSS 02-MAY-2002
DEFINITION  SALK_064078 Arabidopsis thaliana TDNA insertion lines Arabidopsis
            thaliana genomic clone SALK_064078, DNA sequence.
ACCESSION   BH813384
VERSION     BH813384.1  GI:20391857
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within 300 bases of the 5' end of
            At1g54490.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_064078"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       26 a      4 c      17 g      12 t
ORIGIN

  Query Match             1.0%;  Score 23.2;  DB 17;  Length 59;
  Best Local Similarity   65.4%;  Pred. No. 1.1e+06;
  Matches   34;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy    2163 TGAAAAACATCCCAAACTTTAAAATGCGAAATTATTGGTTGGTGTGAAGAAA 2214
           | ||||| | | || |  ||||   |  |||| | |  |||| ||||||| |
Db       4 TAAAAAAGAACTCAGAGGTTAATCGGAAAAATAAATATTTGGGGTGAAGAGA 55


RESULT 11
BG271427
LOCUS       BG271427                  50 bp    mRNA    linear   EST 20-FEB-2001
DEFINITION  nai50f05.x1 NCI_CGAP_HN20 Homo sapiens cDNA clone IMAGE:4263681 3',
            mRNA sequence.
ACCESSION   BG271427
VERSION     BG271427.1  GI:12979558
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   NCI/NIDR-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Dental Research,
```

```
            Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4263681"
                     /clone_lib="NCI_CGAP_HN20"
                     /lab_host="DH10B"
                     /note="Organ: normal head/neck tissue; Vector: pAMP1; mRNA
                     made from head/neck tissue, cDNA made by oligo-dT
                     priming. Directionally cloned into UDG sites.
                     Size-selected on  agarose gel, average insert size 300 bp.
                     Primary library.  cDNA Library Preparation: David B.
                     Krizman, Ph.D."
BASE COUNT       13 a      7 c      9 g     21 t
ORIGIN

  Query Match             1.0%;  Score 23;  DB 12;  Length 50;
  Best Local Similarity   68.1%;  Pred. No. 1.2e+06;
  Matches   32;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy     203 TTGGACCATATGAACTTGGCATGGAACATTGTGAGAAATTTGAAATC 249
           ||   || ||||| | ||  || | |||| || ||||  ||||| ||
Db       2 TTTTTCCTTATGAGCATGTAATTGTACATCGTTAGAAGCTTGAATTC 48


RESULT 12
AZ508445
LOCUS       AZ508445                 51 bp    DNA     linear   GSS 05-OCT-2000
DEFINITION  1M0350F11R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0350F11 R, DNA sequence.
ACCESSION   AZ508445
VERSION     AZ508445.1  GI:10689677
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0350  row: F  column: 11
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 51.
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0350F11"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       11 a      4 c      0 g     36 t
ORIGIN

  Query Match             1.0%;  Score 23;  DB 17;  Length 51;
  Best Local Similarity   68.1%;  Pred. No. 1.2e+06;
  Matches   32;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy    2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCA 2154
           |||||| ||||| ||  ||| ||||| |  | |   || || ||| |
Db       4 TCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTCATTTA 50


RESULT 13
AZ621827
LOCUS       AZ621827                 59 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0455C10F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0455C10 F, DNA sequence.
ACCESSION   AZ621827
VERSION     AZ621827.1  GI:11743933
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0455  row: C  column: 10
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 59.
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
```

/db_xref="taxon:10090"
/clone="UUGC1M0455C10"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 23 a 7 c 11 g 18 t
ORIGIN

Query Match 1.0%; Score 23; DB 17; Length 59;
Best Local Similarity 63.6%; Pred. No. 1.2e+06;
Matches 35; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 2240 TGGTGAAATAATAAAATGCAAATGAATCATTGTTAACACAGCTGTGGCTCGTTTG 2294
Db 2 TTGATAACTAGTAAGATAGCAAGCAATGATTGTTTACACAAGTGTACATCATATG 56

RESULT 14
AA554929/c
LOCUS AA554929 60 bp mRNA linear EST 08-SEP-1997
DEFINITION nk31g05.s1 NCI_CGAP_Co11 Homo sapiens cDNA clone IMAGE:1015160 3' similar to SW:RL37_HUMAN P02403 60S RIBOSOMAL PROTEIN L37. ;, mRNA sequence.
ACCESSION AA554929
VERSION AA554929.1 GI:2325468
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 60)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Elias Campo, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 481 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham.
FEATURES Location/Qualifiers
source 1. .60
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1015160"
/clone_lib="NCI_CGAP_Co11"
/tissue_type="tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: colon; Vector: Bluescript SK-; Site_1: EcoRI ; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Multiple colon tumors. 5' adaptor sequence: 5' GAATTCGGCACGAG 3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3' Average insert size: 1.1 kb."

BASE COUNT 10 a 3 c 2 g 45 t
ORIGIN

Query Match 1.0%; Score 23; DB 9; Length 60;
Best Local Similarity 68.1%; Pred. No. 1.2e+06;
Matches 32; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1754 GATTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAA 1800
Db 47 GGTTTTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1

RESULT 15
AW411378
LOCUS AW411378 60 bp mRNA linear EST 29-JUN-2000
DEFINITION fh12a03.y1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2964389 3', mRNA sequence.
ACCESSION AW411378
VERSION AW411378.1 GI:6936919
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 60)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Plate: LLCM64 row: B column: 6
Seq primer: M13RP1 reverse primer (ABI).
FEATURES Location/Qualifiers
source 1. .60
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2964389"
/clone_lib="NIH_MGC_17"
/tissue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 18 a 4 c 5 g 33 t
ORIGIN

Query Match 1.0%; Score 23; DB 10; Length 60;
Best Local Similarity 68.1%; Pred. No. 1.2e+06;
Matches 32; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCA 2154
Db 12 TTTTTTTTGTAGAAAAGAATATTTTATTGAAACAGTTTCTCAATTAA 58

```
RESULT 16
BG409218
LOCUS       BG409218                 60 bp    mRNA    linear   EST 13-MAR-2001
DEFINITION  gb88b10.y1 Moss EST library PPG Physcomitrella patens cDNA clone
            PEP_SOURCE_ID:PPG_CopyA-120419 5', mRNA sequence.
ACCESSION   BG409218
VERSION     BG409218.1  GI:13315563
KEYWORDS    EST.
SOURCE      Physcomitrella patens.
  ORGANISM  Physcomitrella patens
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
            Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Quatrano,R., Bashiardes,S., Cove,D., Cuming,A., Knight,C., Clifton
            ,S., Marra,M., Hillier,L., Pape,D., Martin,J., Wylie,T., Underwood
            ,K., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T.,
            Steptoe,M., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.
            , Waterston,R. and Wilson,R.
  TITLE     Leeds/Wash U Moss EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Ralph Quatrano
            Leeds/Wash U Moss EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Libraries were constructed by Dr. Stavros Bashiardes as part of the
            Physcomitrella EST program (PEP) at the Univ. of Leeds (UK) and
            Washington Univ. in St. Louis (USA) DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Celia Knight (c.d.knight@leeds.ac.uk)
            Putative full length read
            vector to vector length is 61
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Physcomitrella patens"
                     /db_xref="taxon:3218"
                     /clone="PEP_SOURCE_ID:PPG_CopyA-120419"
                     /clone_lib="Moss EST library PPG"
                     /tissue_type="gametophore: 30 day old tissue,
                     ammonium-grown"
                     /lab_host="DH10B"
                     /note="Vector: pAMP1; Construction of the cDNA library was
                     performed by Dr. W. Gregg Clark using a modification of
                     the cDNA synthesis protocol developed in the laboratory of
                     Dr. Michael Lovett by Dr. Yulia Korshunova (personal
                     communication). First polyA + RNA was isolated from total
                     gametophore RNA using oligo dT magnetic beads. Following
                     this, first strand cDNA synthesis was performed on the
                     bead-bound polyA + RNA, during which an oligonucleotide
                     anchor sequence was incorporated onto the 5'-ends of the
                     cDNA. PCR amplification was then used to synthesize the
                     second strand, to amplify the double stranded DNA, and to
                     incorporate dUTP containing sequences into the ends of the
                     double stranded cDNA. This DNA was size selected and
                     cloned into pAMP1 using the CloneAMP pAMP1 System (Life
                     Technologies, GibcoBRL) for cloning amplification products
                     by a non-restriction site dependant process. The cloning
                     was  directional based on sequence asymmetry introduced at
                     the ends during PCR amplification. The 3' cDNA ends are
                     proximal to the NotI site of the  multiple cloning site in
                     pAMP1. This annealing mixture was transformed into
                     chemically competent DH10B cells and selected for
                     ampicillin resistant  growth. The resulting clones (about
                     330,000) were pooled to make the library."
BASE COUNT       34 a      2 c      11 g      13 t
ORIGIN

  Query Match             1.0%;  Score 23;  DB 12;  Length 60;
  Best Local Similarity   68.1%;  Pred. No. 1.2e+06;
  Matches   32;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   1742 AAAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAA 1788
           ||| |||  | | |  ||   |||| ||| ||| ||  |||||||||
Db     14 AAAGTAAGTGAGTAAGTTGGTAAATAAATAAATAGTAAAAAAAAAAA 60


RESULT 17
AW249993/c
LOCUS       AW249993                 52 bp    mRNA    linear   EST 07-JAN-2000
DEFINITION  2819345.3prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2819345 3',
            mRNA sequence.
ACCESSION   AW249993
VERSION     AW249993.1  GI:6592986
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: 2819345.5prime
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling
            Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
            Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
            project Clone distribution: MGC clone distribution information can
            be found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
            Scores: PHRED from University of Washington Genome Center. Vector
            Trimming: cross_match from University of Washington Genome Center
            PHRAP suite. Poly-T Identification: patMatch.pl from Berkeley
            Drosophila Genome Project. University of Washington Genome Center:
            http://www.genome.washington.edu Low Quality Sequence: 8 contiguous
            PHRED high quality bases following vector sequence. Very Low
            Quality Sequence: Trace file contained 52 contiguous distinct peaks
            following vector sequence. Polyadenylation: Based upon the presence
            of a XhoI site followed by a run of 14 or more T residues at the
            beginning of the sequence, this cDNA insert was polyadenylated.
            Plate: LLCM1  row: F  column: 18
            High quality sequence stop: 8.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2819345"
                     /clone_lib="NIH_MGC_7"
                     /tissue_type="small cell carcinoma"
                     /cell_line="MGC3"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT        7 a      3 c      6 g      36 t
ORIGIN

  Query Match             1.0%;  Score 22.8;  DB 10;  Length 52;
  Best Local Similarity   66.0%;  Pred. No. 1.3e+06;
  Matches   33;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   1738 AGGAAAAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAA 1787
           ||||||||||||  ||       | |||||  ||  ||   |  ||||||||
Db     50 AGGAAAAATAAAAGTCCTCCCTTAAAAAAAAAAAAAAAACTAAAAAAAAAA 1
```

```
RESULT 18
BM880091/c
LOCUS       BM880091                 53 bp    mRNA    linear   EST 07-MAR-2002
DEFINITION  ku04h02.y1 Strongyloides ratti PA female naive pAMP1 v1
            Strongyloides ratti cDNA 5', mRNA sequence.
ACCESSION   BM880091
VERSION     BM880091.1  GI:19252758
KEYWORDS    EST.
SOURCE      Strongyloides ratti.
  ORGANISM  Strongyloides ratti
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Panagrolaimoidea; Strongyloididae; Strongyloides.
REFERENCE   1  (bases 1 to 53)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.
            , Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.
            , Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R.
            , Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
            ,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.
            , Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
            Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Brandi Chiapelli and Dr. James
            McCarter at Washington University, St. Louis. The cDNA was made by
            using Dynabead oligo-dT priming (Dynal). PCR based library using a
            modified protocol from the SMART PCR cDNA Synthesis Kit from
            Clontech. Directionally cloned into the UDG sites of pAMP1.
            Dissected nematode tissues were provided by Dr. Alan Scott
            (ascott@jhsph.edu) of the School of Public Hygene and Public Health
            at John Hopkins University in Baltimore, MD.
FEATURES             Location/Qualifiers
     source          1. .53
                     /organism="Strongyloides ratti"
                     /db_xref="taxon:34506"
                     /clone_lib="Strongyloides ratti PA female naive pAMP1 v1"
                     /dev_stage="parasitic adult females"
                     /lab_host="DH10B"
                     /note="Vector: pAMP1 (Gibco); Site_1: NotI; Site_2: SalI;
                     The library was constructed by Claire Murphy, Brandi
                     Chiapelli and Dr. James McCarter at Washington University,
                     St. Louis. The cDNA was made by using Dynabead oligo-dT
                     priming (Dynal). PCR based library using a modified
                     protocol from the SMART PCR cDNA Synthesis Kit from
                     Clontech. Directionally cloned into the UDG sites of
                     pAMP1. Parasitic adult females were collected from naive
                     animals and provided by Dr. Mark Viney of Bristol, UK."
BASE COUNT       36 a      2 c      2 g     13 t
ORIGIN

  Query Match             1.0%;  Score 22.8;  DB 14;  Length 53;
  Best Local Similarity   71.4%;  Pred. No. 1.3e+06;
  Matches   30;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTG 2149
           | |||||||| || |||||| | ||||   |||   || |||
Db     47 TTTTTTTTTTTTAAAAATATTTTTTTTATTAATTATTTATTG 6


RESULT 19
AZ482068/c
LOCUS       AZ482068                 54 bp    DNA     linear   GSS 04-OCT-2000
DEFINITION  1M0306M23R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0306M23 R, DNA sequence.
ACCESSION   AZ482068
VERSION     AZ482068.1  GI:10643133
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0306  row: M  column: 23
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 54.
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0306M23"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       31 a      4 c      9 g     10 t
ORIGIN

  Query Match             1.0%;  Score 22.8;  DB 17;  Length 54;
  Best Local Similarity   71.4%;  Pred. No. 1.3e+06;
  Matches   30;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   2094 GCAAACAACCTGAATCTTTTTTTTATATAAATATATATTTTT 2135
           |||||||  |  ||||||||||||| | |    | | | |||||
Db     42 GCAAACACTCGAAATCTTTTTTTTTTTTTTTTTTTTTTTTT 1


RESULT 20
AU264598/c
LOCUS       AU264598                 59 bp    mRNA    linear   EST 10-MAY-2002
DEFINITION  AU264598 VS Dictyostelium discoideum cDNA clone VSD814 5', mRNA
            sequence.
ACCESSION   AU264598
VERSION     AU264598.1  GI:20523396
KEYWORDS    EST.
```

```
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Urushihara,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
            Takeuchi,I., Kohara,Y. and Tanaka,Y.
  TITLE     Population analysis of cDNAs from unicellular and multicellular
            stages of Dictyostelium discoideum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="VSD814"
                     /clone_lib="VS"
                     /sex="mat A"
                     /dev_stage="vegetative"
BASE COUNT       30 a      4 c      1 g     24 t
ORIGIN

  Query Match             1.0%;  Score 22.8;  DB 9;  Length 59;
  Best Local Similarity   62.1%;  Pred. No. 1.3e+06;
  Matches   36;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy   2110 TTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCAGCTCATTATGAAA 2167
          ||||||||  ||  ||  |  ||  |||   |  |  |||||  |  |  |  ||  |  ||||
Db     59 TTTTTTTTTTAATAAAAGTTAATTTAATACAATGTTTTTAAATTTGATTATAAAGAAA 2


RESULT 21
AU267788/c
LOCUS       AU267788                  59 bp    mRNA    linear   EST 10-MAY-2002
DEFINITION  AU267788 VS Dictyostelium discoideum cDNA clone VSH663 5', mRNA
            sequence.
ACCESSION   AU267788
VERSION     AU267788.1  GI:20526586
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Urushihara,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
            Takeuchi,I., Kohara,Y. and Tanaka,Y.
  TITLE     Population analysis of cDNAs from unicellular and multicellular
            stages of Dictyostelium discoideum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="VSH663"
                     /clone_lib="VS"
                     /sex="mat A"
                     /dev_stage="vegetative"
BASE COUNT       30 a      4 c      1 g     24 t
ORIGIN

  Query Match             1.0%;  Score 22.8;  DB 9;  Length 59;
  Best Local Similarity   62.1%;  Pred. No. 1.3e+06;
  Matches   36;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy   2110 TTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCAGCTCATTATGAAA 2167
          ||||||||  ||  ||  |  ||  |||   |  |  |||||  |  |  |  ||  |  ||||
Db     59 TTTTTTTTTTAATAAAAGTTAATTTAATACAATGTTTTTAAATTTGATTATAAAGAAA 2


RESULT 22
BQ397088/c
LOCUS       BQ397088                  59 bp    mRNA    linear   EST 22-MAY-2002
DEFINITION  NISC_ng25f10.x1 NICHD XGC Emb6 Silurana tropicalis cDNA clone
            IMAGE:5384586 3', mRNA sequence.
ACCESSION   BQ397088
VERSION     BQ397088.1  GI:21084775
KEYWORDS    EST.
SOURCE      western clawed frog.
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   NIH-XCG http://image.llnl.gov/image/html/xenopuslib_info.shtml.
  TITLE     National Institute of Child Health and Human Development, National
            Cancer Insitute, Xenopus Gene Collection
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation:
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
             DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC)
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            info@image.llnl.gov
            Plate: LLAM11979  row: K  column: 19
            Seq primer: -21M13 forward primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Silurana tropicalis"
                     /db_xref="taxon:8364"
                     /clone="IMAGE:5384586"
                     /clone_lib="NICHD XGC Emb6"
                     /tissue_type="neurula"
                     /dev_stage="embryo, stages 14-19"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Vector: pCMV-SPORT6.ccdb; Site_1: NotI; Site_2:
                     EcoRV; Cloned unidirectionally. Primer: Oligo dT. Average
                     insert size 2.1 kb. Constructed by Invitrogen. Note: This
                     is a Xenopus Gene Collection (XGC) library."
BASE COUNT        2 a      2 c      1 g     54 t
ORIGIN

  Query Match             1.0%;  Score 22.8;  DB 14;  Length 59;
  Best Local Similarity   62.1%;  Pred. No. 1.3e+06;
  Matches   36;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy   1743 AAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAA 1800
          |||  ||||    |   ||||||||  ||  ||      |||||||||   |||  ||||
Db     59 AAAAAAACAAAAAAAGTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2


RESULT 23
AU266759/c
LOCUS       AU266759                  51 bp    mRNA    linear   EST 10-MAY-2002
DEFINITION  AU266759 VS Dictyostelium discoideum cDNA clone VSG728 5', mRNA
            sequence.
ACCESSION   AU266759
VERSION     AU266759.1  GI:20525557
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
```

```
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Urushihara,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
            Takeuchi,I., Kohara,Y. and Tanaka,Y.
  TITLE     Population analysis of cDNAs from unicellular and multicellular
            stages of Dictyostelium discoideum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="VSG728"
                     /clone_lib="VS"
                     /sex="mat A"
                     /dev_stage="vegetative"
BASE COUNT       30 a      3 c      3 g     15 t
ORIGIN

  Query Match             1.0%;  Score 22.6;  DB 9;  Length 51;
  Best Local Similarity   68.9%;  Pred. No. 1.5e+06;
  Matches   31;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
          | |||||| |||||  | |||||| | |  |  |  | |||||||
Db     49 TTTTTTTTCTATATTTAAATATATCTATATACAAATTATTTGATT 5


RESULT 24
AU006675/c
LOCUS       AU006675                  53 bp    mRNA    linear   EST 31-JUL-1998
DEFINITION  AU006675 Schizosaccharomyces pombe late log phase cDNA
            Schizosaccharomyces pombe cDNA clone spc00272, mRNA sequence.
ACCESSION   AU006675
VERSION     AU006675.1  GI:3343134
KEYWORDS    EST.
SOURCE      fission yeast.
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 53)
  AUTHORS   Morimyo,M. and Mita,K.
  TITLE     Identification of expressed sequence tags of Schizosaccharomyces
            pombe
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mitsuoki Morimyo
            Genome Research Group
            National Institute of Radiological Sciences
            9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
            Email: morimyo@nirs.go.jp.
FEATURES             Location/Qualifiers
     source          1. .53
                     /organism="Schizosaccharomyces pombe"
                     /strain="972"
                     /db_xref="taxon:4896"
                     /clone="spc00272"
                     /clone_lib="Schizosaccharomyces pombe late log phase cDNA"
                     /sex="h minus"
                     /note="Vector: M13mp19; The cDNA library of
                     Schizosaccharomyces pombe was prepared by cloning cDNA
                     into the SmaI site of M13mp19 DNA and the direction of DNA
                     sequences was not always from 5' to 3'. The cDNA data of
                     Schizosaccharomyces pombe are available for searching on
                     the World Wide Web.  (URL, http://www.nirs.go.jp)"
BASE COUNT       23 a      6 c      5 g     19 t
ORIGIN

  Query Match             1.0%;  Score 22.6;  DB 9;  Length 53;
  Best Local Similarity   75.7%;  Pred. No. 1.5e+06;
  Matches   28;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy   2110 TTTTTTTTATATAAATATATATTTTTCAAATAGATTT 2146
           ||||||| ||| | |||  | ||||||||||   ||
Db     53 TTTTTTTAATAGATATAGGTCTTTTTCAAATATCATT 17


RESULT 25
AU264938
LOCUS       AU264938                  56 bp    mRNA    linear   EST 10-MAY-2002
DEFINITION  AU264938 VS Dictyostelium discoideum cDNA clone VSF251 5', mRNA
            sequence.
ACCESSION   AU264938
VERSION     AU264938.1  GI:20523736
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   Urushihara,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
            Takeuchi,I., Kohara,Y. and Tanaka,Y.
  TITLE     Population analysis of cDNAs from unicellular and multicellular
            stages of Dictyostelium discoideum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .56
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="VSF251"
                     /clone_lib="VS"
                     /sex="mat A"
                     /dev_stage="vegetative"
BASE COUNT       36 a      2 c      0 g     16 t      2 others
ORIGIN

  Query Match             1.0%;  Score 22.6;  DB 9;  Length 56;
  Best Local Similarity   66.0%;  Pred. No. 1.5e+06;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   1742 AAAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAA 1788
          |||||||||     ||| ||| ||  || | ||  |   ||||||||
Db      8 AAAATAAACNAAATTTTAAAATAAAAAAACNATTTTTTTAAAAAAAA 54


RESULT 26
B03428
LOCUS       B03428                    56 bp    DNA     linear   GSS 13-JUL-1996
DEFINITION  cSRL-179A8-u cSRL flow sorted Chromosome 11 specific cosmid Homo
            sapiens genomic clone cSRL-179A8, DNA sequence.
ACCESSION   B03428
VERSION     B03428.1  GI:1412706
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   Evans,G.A., Burbee,D., Davies,C., Hahner,L., Oliver,T., Gilbert,M.,
            Jones,D., Ward,T., Gillilan,E., Schagemann,J., Probst,S. ., Harris
            ,J., DeFord,J., McFarland,J., Burzinski,K., Khan,M., Kupfer,K. and
            Garner,H.R.
```

```
  TITLE     Genomic Sequence Sampled Map of Chromosome 11
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Evans GA, Shane Probst
            McDermott Center for Human Growth and Development
            University of Texas Southwestern Medical Center At Dallas
            5323 Harry Hines Blvd, Dallas TX 75235-8591
            Tel: 214-648-1600
            Fax: 214-648-1666
            Email: gevans@utsw.swmed.edu, shane@mcdermott.swmed.edu
            Seq primer: T7
            Class: cosmid ends
            High quality sequence stop: 56.
FEATURES             Location/Qualifiers
     source          1. .56
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="cSRL-179A8"
                     /clone_lib="cSRL flow sorted Chromosome 11 specific
                     cosmid"
                     /sex="female"
                     /cell_type="chimeric hamster somatic cell hybrid"
                     /note="Vector: sCos-1; Human Chromosome 11 specific cosmid
                     library prepared from flow sorted human  Chromosome 11
                     derived from Chinese Hampster Ovary (CHO) monochromosomal
                     somatic cell hybrid, J1"
BASE COUNT       24 a      4 c      1 g     26 t      1 others
ORIGIN

  Query Match             1.0%;  Score 22.6;  DB 17;  Length 56;
  Best Local Similarity   73.7%;  Pred. No. 1.5e+06;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   2115 TTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
          || | ||||| || | ||||| |||||  ||||  |||
Db     17 TTAAAATAAAAATTTTTTTTTTAAATAANTTTTAAATT 54


RESULT 27
AA815875/c
LOCUS       AA815875                 57 bp    mRNA    linear   EST 13-FEB-1998
DEFINITION  vp74b01.r1 Knowles Solter mouse blastocyst B3 Mus musculus cDNA
            clone IMAGE:1082377 5', mRNA sequence.
ACCESSION   AA815875
VERSION     AA815875.1  GI:2885471
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 57)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:596673.
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="Mus musculus"
                     /strain="C57BL/6J x DBA/2J F1"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1082377"
                     /clone_lib="Knowles Solter mouse blastocyst B3"
                     /tissue_type="blastocyst"
                     /dev_stage="embryo (pre-implantation)"
                     /lab_host="DH10B"
                     /note="Organ: embryo; Vector: pSPORT; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally from mRNA prepared
                     from 800 blastocysts.  Primer: SalI(dT):
                     5'-CGGTCGACCGTCGACCGTTTTTTTTTTTTTT-3'.  cDNAs  were
                     cloned into the NotI/SalI sites of a pSPORT vector (Life
                     Technologies).  Two different size selections:  B1 (larger
                     inserts) and B3."
BASE COUNT       50 a      4 c      1 g      2 t
ORIGIN

  Query Match             1.0%;  Score 22.6;  DB 9;  Length 57;
  Best Local Similarity   68.9%;  Pred. No. 1.5e+06;
  Matches   31;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
          | ||||||||| | |   | | | |||||||  | | ||||||
Db     53 TTTTTTTTTTTTTTTTTTTTTTTTTTTCATGTTGTTTTTTGTTT 9


RESULT 28
B00931
LOCUS       B00931                   58 bp    DNA     linear   GSS 13-JUL-1996
DEFINITION  cSRL-122a1-u cSRL flow sorted Chromosome 11 specific cosmid Homo
            sapiens genomic clone cSRL-122a1, DNA sequence.
ACCESSION   B00931
VERSION     B00931.1  GI:1410209
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Evans,G.A., Burbee,D., Davies,C., Hahner,L., Oliver,T., Gilbert,M.,
            Jones,D., Ward,T., Gillilan,E., Schagemann,J., Probst,S. ., Harris
            ,J., DeFord,J., McFarland,J., Burzinski,K., Khan,M., Kupfer,K. and
            Garner,H.R.
  TITLE     Genomic Sequence Sampled Map of Chromosome 11
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Evans GA, Shane Probst
            McDermott Center for Human Growth and Development
            University of Texas Southwestern Medical Center At Dallas
            5323 Harry Hines Blvd, Dallas TX 75235-8591
            Tel: 214-648-1600
            Fax: 214-648-1666
            Email: gevans@utsw.swmed.edu, shane@mcdermott.swmed.edu
            Seq primer: T7
            Class: cosmid ends
            High quality sequence stop: 58.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="cSRL-122a1"
                     /clone_lib="cSRL flow sorted Chromosome 11 specific
                     cosmid"
                     /sex="female"
                     /cell_type="chimeric hamster somatic cell hybrid"
                     /note="Vector: sCos-1; Human Chromosome 11 specific cosmid
                     library prepared from flow sorted human  Chromosome 11
                     derived from Chinese Hampster Ovary (CHO) monochromosomal
                     somatic cell hybrid, J1"
BASE COUNT       30 a      3 c      5 g     19 t      1 others
ORIGIN

  Query Match             1.0%;  Score 22.6;  DB 17;  Length 58;
  Best Local Similarity   67.4%;  Pred. No. 1.5e+06;
  Matches   31;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2096 AAACAACCTGAATCTTTTTTTTATATAAATATATATTTTTCAAATA 2141
```

```
          | | ||    ||  |||||| |||| |||||| | |||| ||||
Db      5 ANAAAAAAAAAAGGTTTTTTATATAAAAATATTAAATTTTATAATA 50


RESULT 29
CNS06CC9
LOCUS       CNS06CC9                  60 bp    DNA     linear   GSS 17-JUN-2001
DEFINITION  T3 end of clone AR0AA003G11 of library AR0AA from strain CBS 732 of
            Zygosaccharomyces rouxii, genomic survey sequence.
ACCESSION   AL392335
VERSION     AL392335.1  GI:12141583
KEYWORDS    GSS.
SOURCE      Zygosaccharomyces rouxii.
  ORGANISM  Zygosaccharomyces rouxii
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Zygosaccharomyces.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
  MEDLINE   20584711
   PUBMED   11152876
REFERENCE   2  (bases 1 to 60)
  AUTHORS   de Montigny,J., Straub,M., Potier,S., Tekaia,F., Dujon,B.,
            Wincker,P., Artiguenave,F. and Souciet,J.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 8.
            Zygosaccharomyces rouxii
  JOURNAL   FEBS Lett. 487 (1), 52-55 (2000)
  MEDLINE   20584718
   PUBMED   11152883
REFERENCE   3  (bases 1 to 60)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var.  hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Zygosaccharomyces rouxii"
                     /strain="CBS 732"
                     /db_xref="taxon:4956"
                     /clone="AR0AA003G11"
                     /clone_lib="AR0AA"
                     /note="end : T3"
BASE COUNT       5 a      1 c      0 g     53 t      1 others
ORIGIN

  Query Match             1.0%;  Score 22.6;  DB 17;  Length 60;
  Best Local Similarity   66.0%;  Pred. No. 1.5e+06;
  Matches   31;  Conservative    1;  Mismatches   15;  Indels    0;  Gaps    0;

Qy    2106 AATCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
           |:| ||||||||||| |  || ||| |||||    |   |||||  ||
Db      10 AWTTTTTTTTTTATTTTCATTTATTTTTTTTTTTTTTTTTTTTTTT 56


RESULT 30
```

```
AU265345/c
LOCUS       AU265345                  59 bp    mRNA    linear   EST 10-MAY-2002
DEFINITION  AU265345 VS Dictyostelium discoideum cDNA clone VSF512 5', mRNA
            sequence.
ACCESSION   AU265345
VERSION     AU265345.1  GI:20524143
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Urushihara,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
            Takeuchi,I., Kohara,Y. and Tanaka,Y.
  TITLE     Population analysis of cDNAs from unicellular and multicellular
            stages of Dictyostelium discoideum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="VSF512"
                     /clone_lib="VS"
                     /sex="mat A"
                     /dev_stage="vegetative"
BASE COUNT       30 a      0 c      4 g     17 t      8 others
ORIGIN

  Query Match             1.0%;  Score 22.4;  DB 9;  Length 59;
  Best Local Similarity   57.1%;  Pred. No. 1.6e+06;
  Matches   32;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy    2080 TTGATGTTTTACGTGCAAACAACCTGAATCTTTTTTTTATATAAATATATATTTTT 2135
           ||  | |||||      ||  || |     ||||||||  | || |  || |||||
Db      59 TTTTTTTTTAAAANNNAAAAANCCCCNAATTTTTTTTTNAAAANTNAATTTTTTT 4


RESULT 31
AU263817
LOCUS       AU263817                  60 bp    mRNA    linear   EST 10-MAY-2002
DEFINITION  AU263817 VS Dictyostelium discoideum cDNA clone VSD310 3', mRNA
            sequence.
ACCESSION   AU263817
VERSION     AU263817.1  GI:20522615
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Urushihara,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
            Takeuchi,I., Kohara,Y. and Tanaka,Y.
  TITLE     Population analysis of cDNAs from unicellular and multicellular
            stages of Dictyostelium discoideum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
```

```
                     /clone="VSD310"
                     /clone_lib="VS"
                     /sex="mat A"
                     /dev_stage="vegetative"
BASE COUNT      23 a      3 c      1 g      22 t      11 others
ORIGIN

  Query Match             1.0%;  Score 22.4;  DB 9;  Length 60;
  Best Local Similarity   54.2%;  Pred. No. 1.6e+06;
  Matches   32;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy   2062 ACAAAAGAAACTTACCAATTGATGTTTTACGTGCAAACAACCTGAATCTTTTTTTTATA 2120
          | ||||  ||   |    ||| | ||||   |  ||| ||||   |  | | ||| | |
Db      1 AAAAAAATAAACAATTTNTTGTTNTTTTTTTTTNAAAAAACCNNNANNTNTNTTTNAAA 59


RESULT 32
AU268768/c
LOCUS       AU268768                 44 bp    mRNA    linear   EST 10-MAY-2002
DEFINITION  AU268768 VS Dictyostelium discoideum cDNA clone VSI483 5', mRNA
            sequence.
ACCESSION   AU268768
VERSION     AU268768.1  GI:20527566
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Urushihara,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
            Takeuchi,I., Kohara,Y. and Tanaka,Y.
  TITLE     Population analysis of cDNAs from unicellular and multicellular
            stages of Dictyostelium discoideum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="VSI483"
                     /clone_lib="VS"
                     /sex="mat A"
                     /dev_stage="vegetative"
BASE COUNT      30 a      0 c      0 g      14 t
ORIGIN

  Query Match             0.9%;  Score 22.2;  DB 9;  Length 44;
  Best Local Similarity   69.8%;  Pred. No. 1.8e+06;
  Matches   30;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGA 2150
          | ||||||| |||||  |||||||||    |||  || || |
Db     43 TTTTTTTTTAATATATTTATATATTTAATTAATTATTTATTTA 1


RESULT 33
AU268819/c
LOCUS       AU268819                 44 bp    mRNA    linear   EST 10-MAY-2002
DEFINITION  AU268819 VS Dictyostelium discoideum cDNA clone VSI519 3', mRNA
            sequence.
ACCESSION   AU268819
VERSION     AU268819.1  GI:20527617
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Urushihara,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
            Takeuchi,I., Kohara,Y. and Tanaka,Y.
  TITLE     Population analysis of cDNAs from unicellular and multicellular
            stages of Dictyostelium discoideum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="VSI519"
                     /clone_lib="VS"
                     /sex="mat A"
                     /dev_stage="vegetative"
BASE COUNT      30 a      2 c      1 g      10 t      1 others
ORIGIN

  Query Match             0.9%;  Score 22.2;  DB 9;  Length 44;
  Best Local Similarity   68.2%;  Pred. No. 1.8e+06;
  Matches   30;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   2104 TGAATCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTT 2147
          | | | |||||||| ||| ||| | | ||||| || ||    ||
Db     44 TAATTTTTTTTTTTTTATTAATTTTTTTTTTAAACTAAGGNTT 1


RESULT 34
BF643229
LOCUS       BF643229                 52 bp    mRNA    linear   EST 20-DEC-2000
DEFINITION  NF002B08EC1F1062 Elicited cell culture Medicago truncatula cDNA
            clone NF002B08EC 5', mRNA sequence.
ACCESSION   BF643229
VERSION     BF643229.1  GI:11908354
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
            Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation -
            Center for Medicago Genomics Research
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Dixon RA
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7302
            Fax: 580 221 7380
            Email: radixon@noble.org
            Insert Length: 52   Std Error: 0.00
            Plate: 002  row: B  column: 08
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF002B08EC"
                     /clone_lib="Elicited cell culture"
                     /tissue_type="Cell cultures derived from root tissues"
                     /dev_stage="Cell suspensions were subcultured every 14
                     days. Cells were induced six days after subculture"
                     /note="Vector: Lambda Zap; Cells were induced with yeast
```

cell wall extracts equivalent to 50ug/ml glucose in the final concentration. Samples were taken at 0.5, 1, 12 and 24 hours after induction. Equal amounts of RNA from each time point were pooled and used for mRNA isolation."
BASE COUNT 24 a 6 c 2 g 20 t
ORIGIN

Query Match 0.9%; Score 22.2; DB 12; Length 52;
Best Local Similarity 64.7%; Pred. No. 1.8e+06;
Matches 33; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 2114 TTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCAGCTCATTATG 2164
|| ||||| |||| |||| | || || | || || || ||||| |
Db 2 TTCAATATATATATCTATTAATAAATTAAAACATTCATACATTACATTAAG 52

RESULT 35
TA385D03Q/c
LOCUS TA385D03Q 53 bp DNA linear GSS 13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 385d03, reverse sequence, genomic survey sequence.
ACCESSION AL498857
VERSION AL498857.1 GI:11874579
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
REFERENCE 1 (bases 1 to 53)
AUTHORS Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R., Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L., Melville,S.E., Rajandream,M.A. and Barrell,B.G.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nh1@sanger.ac.uk
COMMENT Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (
4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES Location/Qualifiers
source 1. .53
/organism="Trypanosoma brucei"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="385d03"
BASE COUNT 19 a 4 c 8 g 22 t
ORIGIN

Query Match 0.9%; Score 22.2; DB 17; Length 53;
Best Local Similarity 77.1%; Pred. No. 1.8e+06;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2101 ACCTGAATCTTTTTTTTATATAAATATATATTTTT 2135
|| ||||||| | | | |||||| |||||||| | |
Db 42 ACATGAATCATATATATATATATATATATATATAT 8

RESULT 36
AA937418/c
LOCUS AA937418 54 bp mRNA linear EST 30-APR-1998
DEFINITION oj09a06.s1 NCI_CGAP_Mel3 Homo sapiens cDNA clone IMAGE:1491634 3' similar to gb:X17206 40S RIBOSOMAL PROTEIN S4 (HUMAN);, mRNA sequence.
ACCESSION AA937418
VERSION AA937418.1 GI:3095529
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 54)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
unknown library type
Seq primer: -40m13 fwd. ET from Amersham.
FEATURES Location/Qualifiers
source 1. .54
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1491634"
/clone_lib="NCI_CGAP_Mel3"
/tissue_type="metastatic melanoma to bowel"
/lab_host="DH10B"
/note="Organ: bowel (skin primary); Vector: pCMV-SPORT4; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 0.9 kb. Life Technologies catalog #: 10981-017"
BASE COUNT 5 a 5 c 1 g 43 t
ORIGIN

Query Match 0.9%; Score 22.2; DB 9; Length 54;
Best Local Similarity 69.8%; Pred. No. 1.8e+06;
Matches 30; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1760 AAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
| ||| | ||| || ||| ||||||||| ||| |||| |
Db 50 ATAAAGTGAATTAAGCGTGAAAAAAAAAAAAAAAAAAAAAAAA 8

RESULT 37
AZ760065
LOCUS AZ760065 55 bp DNA linear GSS 16-FEB-2001
DEFINITION 1M0553B02R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0553B02 R, DNA sequence.
ACCESSION AZ760065
VERSION AZ760065.1 GI:12867501
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 55)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0553 row: B column: 02
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 55.

```
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0553B02"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       19 a     11 c     15 g     10 t
ORIGIN

  Query Match             0.9%;  Score 22.2;  DB 17;  Length 55;
  Best Local Similarity   77.1%;  Pred. No. 1.8e+06;
  Matches   27;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    766 GGAACAGTCACACACACATTTTGTGGAACAATAGA 800
          || ||||| |||||||||||| ||| || |   ||
Db     21 GGCACAGTGACACACACATTTGGTGTAAGAGCTGA 55


RESULT 38
BI088641
LOCUS       BI088641                 56 bp    mRNA    linear   EST 20-JUN-2001
DEFINITION  602851403F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4992768 5',
            mRNA sequence.
ACCESSION   BI088641
VERSION     BI088641.1  GI:14506971
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Incyte Genomics, Inc.
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11012  row: f  column: 01
            High quality sequence stop: 56.
FEATURES             Location/Qualifiers
     source          1. .56
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4992768"
                     /clone_lib="NIH_MGC_10"
                     /cell_line="MGC36"
                     /lab_host="DH10B"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.5 kb. Library prepared by Life
                     Technologies."
BASE COUNT       20 a     12 c     16 g      8 t
ORIGIN

  Query Match             0.9%;  Score 22.2;  DB 13;  Length 56;
  Best Local Similarity   69.8%;  Pred. No. 1.8e+06;
  Matches   30;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    398 TGCTTAAAAAGGCAATGATAGTAAGAAATGCTAAAGATACAGC 440
          |||   ||| || ||||| || ||||| | | || |  |||||
Db      6 TGCAGCAAACGGAAATGACAGCAAGAAGTTCAAAGGTGACAGC 48


RESULT 39
AU270417
LOCUS       AU270417                 57 bp    mRNA    linear   EST 10-MAY-2002
DEFINITION  AU270417 VS Dictyostelium discoideum cDNA clone VSJ752 5', mRNA
            sequence.
ACCESSION   AU270417
VERSION     AU270417.1  GI:20529215
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 57)
  AUTHORS   Urushihara,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
            Takeuchi,I., Kohara,Y. and Tanaka,Y.
  TITLE     Population analysis of cDNAs from unicellular and multicellular
            stages of Dictyostelium discoideum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="VSJ752"
                     /clone_lib="VS"
                     /sex="mat A"
                     /dev_stage="vegetative"
BASE COUNT       37 a      2 c      4 g     14 t
ORIGIN

  Query Match             0.9%;  Score 22.2;  DB 9;  Length 57;
  Best Local Similarity   64.7%;  Pred. No. 1.8e+06;
  Matches   33;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy   1738 AGGAAAAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAA 1788
          || ||| || ||   |  ||| |||||| ||  ||   |  |||||||||
Db      7 AGAAAATATTAATAAGTTTTTAAAAAAAAAAAAAAAAATTCAAAAAAAAA 57


RESULT 40
BG943501/c
LOCUS       BG943501                 58 bp    mRNA    linear   EST 11-JUN-2001
DEFINITION  ax38e11.x1 Proliferating Human Erythroid Cells (LCB:ax library)
            Homo sapiens cDNA clone ax38e11 random, mRNA sequence.
ACCESSION   BG943501
VERSION     BG943501.1  GI:14342873
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Gubin,A.N., Njoroge,J.M., Bouffard,G.G. and Miller,J.L.
  TITLE     Gene expression in proliferating human erythroid cells
  JOURNAL   Genomics 59 (2), 168-177 (1999)
  MEDLINE   99339981
COMMENT     Contact: Jeffery L. Miller
            Laboratory of Chemical Biology
            National Institute of Diabetes and Digestive and Kidney Diseases
            Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
            20892, USA
            Tel: 301 402 2373
            Fax: 301 435 5148
            Email: jm7f@nih.gov
            DNA Sequencing and analyses by National Institutes of Health
            Intramural Sequencing Center (NISC).
            Plate: 38  row: e  column: 11
            Seq primer: -21M13 forward primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="ax38e11"
                     /clone_lib="Proliferating Human Erythroid Cells (LCB:ax
                     library)"
                     /sex="unknown"
                     /tissue_type="blood"
                     /cell_type="Erythroid Cells"
                     /cell_line="Primary Culture of Peripheral Blood
                     Mononuclear Cells"
                     /dev_stage="Progenitor; EPO responsive CD71++++"
                     /lab_host="SOLR"
                     /note="Organ: blood; Vector: Lambda ZAP II; Site_1: EcoRI;
                     Site_2: EcoRI; 65,000 proliferating erythroid cells from
                     the buffy coat of a blood donation were obtained by flow
                     cytometric separation after a 5-day culture period in the
                     presence of erythropoietin. Total RNA was purified from
                     the sorted cell population using TRIzol reagent. RNA (0.3
                     ug) was converted into double stranded cDNA using
                     Clontech's CapFinder cDNA Library Construction Kit
                     (Clontech) according to the manufacturer's protocol and
                     cloned into EcoRI digested Lambda Zap II vector
                     (Stratagene). The phage library was amplified once prior
                     to in vivo excision in SOLR cells. Individual colonies
                     were grown, and the cDNA inserts were sequenced in high
                     throughput (NIH intramural sequencing center
                     http://www.nisc.nih.gov/)."
BASE COUNT        0 a      6 c      2 g     50 t
ORIGIN

  Query Match             0.9%;  Score 22.2;  DB 13;  Length 58;
  Best Local Similarity   69.8%;  Pred. No. 1.8e+06;
  Matches   30;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   1760 AAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
           ||||||  ||  ||  |   |||||||||  ||| |||||  |
Db     47 AAAAAAAAAAAAAAAAGAAAAAAAAAAAACAAAAAACAAAAAA 5


RESULT 41
BG943837/c
LOCUS       BG943837                 59 bp    mRNA    linear   EST 11-JUN-2001
DEFINITION  ax42f06.x1 Proliferating Human Erythroid Cells (LCB:ax library)
            Homo sapiens cDNA clone ax42f06 random, mRNA sequence.
ACCESSION   BG943837
VERSION     BG943837.1  GI:14343209
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Gubin,A.N., Njoroge,J.M., Bouffard,G.G. and Miller,J.L.
  TITLE     Gene expression in proliferating human erythroid cells
  JOURNAL   Genomics 59 (2), 168-177 (1999)
  MEDLINE   99339981
COMMENT     Contact: Jeffery L. Miller
            Laboratory of Chemical Biology
            National Institute of Diabetes and Digestive and Kidney Diseases
            Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
            20892, USA
            Tel: 301 402 2373
            Fax: 301 435 5148
            Email: jm7f@nih.gov
            DNA Sequencing and analyses by National Institutes of Health
            Intramural Sequencing Center (NISC).
            Plate: 42  row: f  column: 06
            Seq primer: -21M13 forward primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="ax42f06"
                     /clone_lib="Proliferating Human Erythroid Cells (LCB:ax
                     library)"
                     /sex="unknown"
                     /tissue_type="blood"
                     /cell_type="Erythroid Cells"
                     /cell_line="Primary Culture of Peripheral Blood
                     Mononuclear Cells"
                     /dev_stage="Progenitor; EPO responsive CD71++++"
                     /lab_host="SOLR"
                     /note="Organ: blood; Vector: Lambda ZAP II; Site_1: EcoRI;
                     Site_2: EcoRI; 65,000 proliferating erythroid cells from
                     the buffy coat of a blood donation were obtained by flow
                     cytometric separation after a 5-day culture period in the
                     presence of erythropoietin. Total RNA was purified from
                     the sorted cell population using TRIzol reagent. RNA (0.3
                     ug) was converted into double stranded cDNA using
                     Clontech's CapFinder cDNA Library Construction Kit
                     (Clontech) according to the manufacturer's protocol and
                     cloned into EcoRI digested Lambda Zap II vector
                     (Stratagene). The phage library was amplified once prior
                     to in vivo excision in SOLR cells. Individual colonies
                     were grown, and the cDNA inserts were sequenced in high
                     throughput (NIH intramural sequencing center
                     http://www.nisc.nih.gov/)."
BASE COUNT        0 a      7 c      1 g     51 t
ORIGIN

  Query Match             0.9%;  Score 22.2;  DB 13;  Length 59;
  Best Local Similarity   69.8%;  Pred. No. 1.8e+06;
  Matches   30;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   1760 AAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
           ||||||  ||  ||  |   |||||||||  ||| |||||  |
Db     50 AAAAAAAAAAAAAAAAGAAGCAAAAAAAAAAAAAAAAAAAAAA 8


RESULT 42
AA014672
LOCUS       AA014672                 49 bp    mRNA    linear   EST 21-JAN-1997
DEFINITION  mh30b01.r1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA
            clone IMAGE:443977 5', mRNA sequence.
ACCESSION   AA014672
VERSION     AA014672.1  GI:1475759
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
```

```
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:269313
            Seq primer: mob.REGA+ET.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:443977"
                     /clone_lib="Soares mouse placenta 4NbMP13.5 14.5"
                     /sex="unknown"
                     /tissue_type="placenta"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: placenta; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       21 a      11 c       3 g      14 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 9;  Length 49;
  Best Local Similarity   67.4%;  Pred. No. 2e+06;
  Matches   31;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy     898 CCCCCATTCACTGGGGAGAATAGAAAGAAAACAATTGACAAAATCC 943
           ||  ||| | ||  |||   ||  || |||| |||| |||||| ||
Db       2 CCTACATCCTCTTTGGAATGTAACAATAAAATAATTTACAAAACCC 47


RESULT 43
BG301974/c
LOCUS       BG301974                 51 bp    mRNA    linear   EST 09-MAY-2001
DEFINITION  kt07h05.y1 Strongyloides ratti L3 adult SL1 Topo1 Chiapelli
            McCarter Strongyloides ratti cDNA 5', mRNA sequence.
ACCESSION   BG301974
VERSION     BG301974.1  GI:13099501
KEYWORDS    EST.
SOURCE      Strongyloides ratti.
  ORGANISM  Strongyloides ratti
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Panagrolaimoidea; Strongyloididae; Strongyloides.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.
            , Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.
            , Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R.
            , Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
            ,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.
            , Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
            Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Brandi Chiapelli and Dr. James
            McCarter (bchiapel@watson.wustl.edu & jmccarte@watson.wustl.edu) at
            Washington University, St. Louis. DNA Sequencing by: Washington
            University Genome Sequencing Center St. Louis.
            Putative full length read
            The vector to vector length is 65.
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Strongyloides ratti"
                     /db_xref="taxon:34506"
                     /clone_lib="Strongyloides ratti L3 adult SL1 Topo1
                     Chiapelli McCarter"
                     /dev_stage="L3/adult"
                     /lab_host="DH10B"
                     /note="Vector: pCRII-TOPO; SL1-Oligo(dT) PCR-based
                     library. Strongyloides_ratti_L3_adult cDNA PCR products of
                     size >400 nucleotides containing SL1 on the 5' end and
                     oligo(dT) on the 3' end were non-directionally cloned into
                     pCRII-TOPO (Invitrogen) following the Topo TA cloning
                     protocol. The cDNA insert can be excised by digestion with
                     EcoRI. The library was constructed by Brandi Chiapelli and
                     Dr. James McCarter at Washington University, St. Louis.
                     Nematodes were provided by Dr. Mark Viney of the
                     University of Bristol, UK."
BASE COUNT       28 a      2 c       5 g      16 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 12;  Length 51;
  Best Local Similarity   73.7%;  Pred. No. 2e+06;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy    2110 TTTTTTTTATATAAATATATATTTTTCAAATAGATTTT 2147
           |||||  ||||  |||  ||| |   ||||||||||||
Db      38 TTTTTAGTATAATAATTAATAATCCACAAATAGATTTT 1


RESULT 44
AI168523/c
LOCUS       AI168523                 55 bp    mRNA    linear   EST 23-OCT-1998
DEFINITION  ox10h11.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA
            clone IMAGE:1655973 3' similar to SW:NU5M_TRYBB P04540
            NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5 ;, mRNA sequence.
ACCESSION   AI168523
VERSION     AI168523.1  GI:3701693
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Insert Length: 349   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1655973"
                     /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
                     /sex="male"
```

```
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     This is a subtracted version of the original Soares fetal
                     liver spleen 1NFLS library.  1st strand cDNA was primed
                     with a Pac I - oligo(dT) primer [5'
                     AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT        5 a      5 c      7 g     38 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 9;  Length 55;
  Best Local Similarity   63.0%;  Pred. No. 2e+06;
  Matches   34;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy   1734 ACACAGGAAAAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAA 1787
           || ||  || || |||  || |    ||||| | ||  ||||  |  |||||||
Db     54 ACCCAAAAACAAGAAAATTGAAAAACAAAAATTAAACAAATGAAGAGAAAAAAA 1


RESULT 45
AU264614
LOCUS       AU264614                  57 bp    mRNA    linear   EST 10-MAY-2002
DEFINITION  AU264614 VS Dictyostelium discoideum cDNA clone VSD826 5', mRNA
            sequence.
ACCESSION   AU264614
VERSION     AU264614.1  GI:20523412
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 57)
  AUTHORS   Urushihara,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
            Takeuchi,I., Kohara,Y. and Tanaka,Y.
  TITLE     Population analysis of cDNAs from unicellular and multicellular
            stages of Dictyostelium discoideum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="VSD826"
                     /clone_lib="VS"
                     /sex="mat A"
                     /dev_stage="vegetative"
BASE COUNT       42 a      1 c      2 g     12 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 9;  Length 57;
  Best Local Similarity   63.0%;  Pred. No. 2e+06;
  Matches   34;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy   1734 ACACAGGAAAAATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAA 1787
           | | |  ||||   ||  |  |  ||||||||| | | |||  |  ||||||||
Db      4 AAAAAAAAAAAAGGCAAAATTAAAATTAAAAAATTATTAAATTTTAAAAAAAAAA 57


RESULT 46
BI745330/c
LOCUS       BI745330                  57 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  rk99b01.y1 Meloidogyne javanica egg pAMP1 v6 Chiapelli McCarter
            Meloidogyne javanica cDNA 5', mRNA sequence.
ACCESSION   BI745330
VERSION     BI745330.1  GI:15767132
KEYWORDS    EST.
SOURCE      root-knot nematode.
  ORGANISM  Meloidogyne javanica
            Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
            Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
REFERENCE   1  (bases 1 to 57)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.
            , Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.
            , Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R.
            , Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
            ,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.
            , Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
            Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Brandi Chiapelli and Dr. James
            McCarter (bchiapel@watson.wustl.edu & jmccarte@watson.wustl.edu) at
            Washington University, St. Louis. DNA Sequencing by: Washington
            University Genome Sequencing Center St. Louis.
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="Meloidogyne javanica"
                     /db_xref="taxon:6303"
                     /clone_lib="Meloidogyne javanica egg pAMP1 v6 Chiapelli
                     McCarter"
                     /dev_stage="enriched for eggs"
                     /lab_host="DH10B"
                     /note="Vector: pAMP1 (Gibco); The library was constructed
                     by Brandi Chiapelli and Dr. James  McCarter at Washington
                     University, St. Louis. The cDNA was made by using Dynabead
                     oligo-dT priming (Dynal). PCR based library using a
                     modified  protocol from the SMART PCR cDNA Synthesis Kit
                     from Clontech. Directionally  cloned into the UDG sites of
                     pAMP1. Nematodes were provided by Dr. David  Bird of North
                     Carolina State University."
BASE COUNT       26 a      1 c      4 g     26 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 13;  Length 57;
  Best Local Similarity   63.0%;  Pred. No. 2e+06;
  Matches   34;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy   1749 ACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
           || |   ||||     | ||||||  |     |||||||  |  |||| |||||||
Db     56 ACTTTTTTTTTTTTTTAACAATCATTTTATTAAAAAAATAAATAAATCAAAATA 3


RESULT 47
BI842365/c
LOCUS       BI842365                  58 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  fq82g06.y1 Zebrafish neuronal Danio rerio cDNA clone 4886962 5',
            mRNA sequence.
ACCESSION   BI842365
VERSION     BI842365.1  GI:15954888
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
```

```
          ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
          ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
          Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
          Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
          and Wilson,R.
  TITLE   WashU Zebrafish EST Project 1998
  JOURNAL Unpublished (1998)
COMMENT   Contact: Stephen L. Johnson
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: zbrafish@watson.wustl.edu
          cDNA Library constructed by S. Lin DNA Sequencing by: Washington
          University Genome Sequencing Center Clone distribution: the
          I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
          Seq primer: T3 ET from Amersham.
FEATURES            Location/Qualifiers
     source         1. .58
                    /organism="Danio rerio"
                    /db_xref="taxon:7955"
                    /clone="4886962"
                    /clone_lib="Zebrafish neuronal"
                    /sex="mixed"
                    /dev_stage="adult"
                    /lab_host="DH10B (phage-resistant)"
                    /note="Organ: brain; Vector: pBluescript (modified);
                    Site_1: DraIII(Y); Site_2: DRAIII(X); Library is cloned
                    directionally between the DraIII(X) and DraIII(Y) sites
                    and has been amplified. Library constructed by S. Lin. "
BASE COUNT      38 a      3 c      7 g     10 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 13;  Length 58;
  Best Local Similarity   67.4%;  Pred. No. 2e+06;
  Matches   31;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTC 2153
           | ||||||||| | |   | | |   |||||||||  ||||| || |
Db     47 TTTTTTTTTTTTTTTTTTTTTTTAATTTCAAATATTTTTTTTATCC 2


RESULT 48
BM965520
LOCUS       BM965520                 58 bp    mRNA    linear   EST 20-MAR-2002
DEFINITION  ko13a10.y1 Toxocara canis adult pAMP1 v1 Toxocara canis cDNA 5',
            mRNA sequence.
ACCESSION   BM965520
VERSION     BM965520.1  GI:19557835
KEYWORDS    EST.
SOURCE      Toxocara canis.
  ORGANISM  Toxocara canis
            Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
            ; Toxocaridae; Toxocara.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.
            , Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.
            , Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R.
            , Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
            ,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.
            , Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
            Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Claire Murphy and Dr. James McCarter
            at Washington University, St. Louis. DNA Sequencing by: Washington
            University Genome Sequencing Center.
FEATURES            Location/Qualifiers
     source         1. .58
                    /organism="Toxocara canis"
                    /db_xref="taxon:6265"
                    /clone_lib="Toxocara canis adult pAMP1 v1"
                    /dev_stage="adult"
                    /lab_host="DH10B"
                    /note="Vector: pAMP1 (Gibco); Site_1: NotI; Site_2: SalI;
                    The library was constructed by Claire Murphy and Dr. James
                    McCarter at Washington University, St. Louis. The cDNA was
                    made by using Dynabead oligo-dT priming (Dynal). PCR based
                    library using a modified protocol from the SMART PCR cDNA
                    Synthesis Kit from Clontech. Directionally cloned into the
                    UDG sites of pAMP1. Adult nematodes were collected from
                    infected dogs and provided by Dr. Prema Arasu of North
                    Carolina State University, Raleigh,NC
                    (Prema_Arasu@ncsu.edu)."
BASE COUNT      29 a      7 c      5 g     16 t      1 others
ORIGIN

  Query Match             0.9%;  Score 22;  DB 14;  Length 58;
  Best Local Similarity   61.8%;  Pred. No. 2e+06;
  Matches   34;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy   1756 TTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATAGTATTGCT 1810
           |||||| ||  ||  |||   | |||||||||   |||| |   |    ||| |
Db      1 TTTTAAATAAATAAAAAATTCCGTAAAAAAAAAAAAAAAGTACTCTGCNGTTGTT 55


RESULT 49
AL762620
LOCUS       AL762620                 58 bp    DNA     linear   GSS 18-JUN-2002
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-026E08-013759,
            genomic survey sequence.
ACCESSION   AL762620
VERSION     AL762620.1  GI:21509009
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 58)
  AUTHORS   Rosso,M., Strizhov,N., Li,Y. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion within the locus defined by clone F4D11. The
            sequences are generated at the MPI for Plant Breeding Research in
            the context of the GABI-Kat project. GABI-Kat is part of the German
            Plant Genomics program designated 'GABI'. Information on line
            availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES            Location/Qualifiers
     source         1. .58
                    /organism="Arabidopsis thaliana"
                    /strain="Columbia 0"
                    /db_xref="taxon:3702"
                    /clone="GK-026E08-013759"
```

/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pAC106. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thalaina nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"
BASE COUNT 23 a 10 c 10 g 15 t
ORIGIN

Query Match 0.9%; Score 22; DB 17; Length 58;
Best Local Similarity 63.0%; Pred. No. 2e+06;
Matches 34; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```
Qy    903 ATTCACTGGGGAGAATAGAAAGAAAACAATTGACAAAATCCTCAAATGTAAACT 956
          ||||| | |  |  |||| || ||| ||  |  |||||  | ||| ||  | ||
Db      4 ATTCATTCGTTAATATAGCAAAAAAGCAGGTTGCAAAACTCGCAATTGATACCT 57
```

RESULT 50
AL779773
LOCUS AL779773 60 bp mRNA linear EST 25-JUN-2002
DEFINITION AL779773 XGC-neurula Silurana tropicalis cDNA clone TNeu078a06 5', mRNA sequence.
ACCESSION AL779773
VERSION AL779773.1 GI:21565477
KEYWORDS EST.
SOURCE western clawed frog.
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Silurana.
REFERENCE 1 (bases 1 to 60)
AUTHORS Taylor,R., Ashurst,J.L., Croning,M.D.R., Zorn,A.M. and Rogers,J.
TITLE Sanger Xenopus tropicalis EST project 2002
JOURNAL Unpublished (2001)
COMMENT Contact: Taylor R
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TNeu078a06.p1cSP6
Sequencing primer: P1CSP6
This sequence is from a Xenopus Gene Collection (XGC) library constructed by Aaron M. Zorn.
FEATURES Location/Qualifiers
source 1. .60
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="TNeu078a06"
/clone_lib="XGC-neurula"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA was oligo dT primed from 5ug of poly A+ RNA from neurula. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT 41 a 8 c 7 g 4 t
ORIGIN

Query Match 0.9%; Score 22; DB 9; Length 60;
Best Local Similarity 67.4%; Pred. No. 2e+06;
Matches 31; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Qy   1760 AAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATAGTA 1805
          |||||||  ||  ||  |  ||||  ||||  |||  |||||  | ||
Db     14 AAAAAATGAAAAAAAAGAAGAAAACAAAAAAAAAAACAAAAAAATA 59
```

RESULT 51
BJ059311
LOCUS BJ059311 38 bp mRNA linear EST 11-DEC-2001
DEFINITION BJ059311 NIBB Mochii normalized Xenopus tailbud library Xenopus laevis cDNA clone XL062d07 5', mRNA sequence.
ACCESSION BJ059311
VERSION BJ059311.1 GI:17490829
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 38)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES Location/Qualifiers
source 1. .38
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL062d07"
/clone_lib="NIBB Mochii normalized Xenopus tailbud library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Nieuwkoop and Faber. Library is subtracted and was constructed by N. Garrett and A.M. Zorn, (Wellcome/CRC Institute). "
BASE COUNT 15 a 1 c 1 g 18 t 3 others
ORIGIN

Query Match 0.9%; Score 21.8; DB 13; Length 38;
Best Local Similarity 72.2%; Pred. No. 2.3e+06;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGA 2143
          | |||||||||| |||||     | ||||| ||| | |
Db      1 TTTTTTTTTTAAATAAANCNGTTTTTTTAAAANAAA 36
```

RESULT 52
BJ057533
LOCUS BJ057533 46 bp mRNA linear EST 10-DEC-2001
DEFINITION BJ057533 NIBB Mochii normalized Xenopus tailbud library Xenopus laevis cDNA clone XL104f09 5', mRNA sequence.
ACCESSION BJ057533
VERSION BJ057533.1 GI:17479603
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 46)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856

```
           Fax: 81-559-81-6855
           Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XL104f09"
                     /clone_lib="NIBB Mochii normalized Xenopus tailbud
                     library"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 25"
                     /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
                     were oligo-dT primed and directionally cloned. Staging
                     according to  Nieuwkoop and Faber. Library is subtracted
                     and was constructed by N. Garrett  and A.M. Zorn,
                     (Wellcome/CRC Institute). "
BASE COUNT        7 a      1 c      7 g     28 t      3 others
ORIGIN

  Query Match             0.9%;  Score 21.8;  DB 13;  Length 46;
  Best Local Similarity   65.9%;  Pred. No. 2.3e+06;
  Matches   29;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2105 GAATCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTT 2148
           || | |||||||| | ||      || ||||| ||    | |||||
Db      3 GAGTTTTTTTTTTTTTTANNGGGATTTTTTTAAATANGTTTTTT 46


RESULT 53
AU263470/c
LOCUS       AU263470                 48 bp    mRNA    linear   EST 10-MAY-2002
DEFINITION  AU263470 VS Dictyostelium discoideum cDNA clone VSD103 5', mRNA
            sequence.
ACCESSION   AU263470
VERSION     AU263470.1  GI:20522268
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Urushihara,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
            Takeuchi,I., Kohara,Y. and Tanaka,Y.
  TITLE     Population analysis of cDNAs from unicellular and multicellular
            stages of Dictyostelium discoideum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="VSD103"
                     /clone_lib="VS"
                     /sex="mat A"
                     /dev_stage="vegetative"
BASE COUNT       38 a      0 c      0 g     10 t
ORIGIN

  Query Match             0.9%;  Score 21.8;  DB 9;  Length 48;
  Best Local Similarity   70.7%;  Pred. No. 2.3e+06;
  Matches   29;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   2112 TTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
          |||||| | ||||| || |||| |    |  |||||| |||
Db     48 TTTTTTTTTTAAATTTAAATTTATTTTTTTTATTTTTTATT 8


RESULT 54
BE545959
LOCUS       BE545959                 48 bp    mRNA    linear   EST 09-AUG-2000
DEFINITION  601069318F2 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:3455548 5',
            mRNA sequence.
ACCESSION   BE545959
VERSION     BE545959.1  GI:9774604
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Incyte Genomics, Inc.
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM8442  row: c  column: 05
            High quality sequence stop: 48.
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3455548"
                     /clone_lib="NIH_MGC_12"
                     /tissue_type="cervical carcinoma cell line"
                     /lab_host="DH10B"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.4 kb. Library prepared by Life
                     Technologies."
BASE COUNT       22 a      3 c      3 g     20 t
ORIGIN

  Query Match             0.9%;  Score 21.8;  DB 10;  Length 48;
  Best Local Similarity   92.0%;  Pred. No. 2.3e+06;
  Matches   23;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy   2106 AATCTTTTTTTTATATAAATATATA 2130
           |||||| |||||||||| |||||||
Db      2 AATCTTATTTTTATATATATATATA 26


RESULT 55
AA920052
LOCUS       AA920052                 49 bp    mRNA    linear   EST 20-APR-1998
DEFINITION  vy63h11.r1 Stratagene mouse macrophage (#937306) Mus musculus cDNA
            clone IMAGE:1310949 5' similar to SW:PSA_HUMAN P55786
            PUROMYCIN-SENSITIVE AMINOPEPTIDASE ;, mRNA sequence.
ACCESSION   AA920052
VERSION     AA920052.1  GI:3066831
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
```

```
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:681245
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1310949"
                     /clone_lib="Stratagene mouse macrophage (#937306)"
                     /tissue_type="macrophage"
                     /dev_stage="WEHI-3 cell line"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: blood; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. WEHI-3 cell line. Average insert size: 1.5 kb;
                     Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT        9 a     12 c     15 g     13 t
ORIGIN

  Query Match             0.9%;  Score 21.8;  DB 9;  Length 49;
  Best Local Similarity   65.3%;  Pred. No. 2.3e+06;
  Matches   32;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy    1903 TTAAGGGTAGGATGTTGCTCTGAATCACTGTGAGTCTGATGTGTGAAGA 1951
           | | ||    |||| |||| |||  |||| |||||||  || ||   ||
Db       1 TGAGGGTCGCGATGCTGCTGTGACACACTTTGAGTCTACTGAGTCCCGA 49


RESULT 56
AA823664
LOCUS       AA823664                 52 bp    mRNA    linear   EST 17-FEB-1998
DEFINITION  vr69d09.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone
            IMAGE:1125905 5' similar to gb:J03161 SERUM RESPONSE FACTOR (HUMAN
            );, mRNA sequence.
ACCESSION   AA823664
VERSION     AA823664.1  GI:2893532
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:615241.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Mus musculus"
                     /strain="C57BL/6J x DBA/2J F1"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1125905"
                     /clone_lib="Knowles Solter mouse 2 cell"
                     /tissue_type="embryo"
                     /dev_stage="2-cell"
                     /lab_host="DH10B"
                     /note="Organ: embryo; Vector: pBluescribe (modified);
                     Site_1: MluI; Site_2: SalI; Cloned unidirectionally from
                     mRNA prepared from 13,500 2-cell stage  embryos.  Primer:
                     SalI(dT): 5'-CGGTCGACCGTCGACCGTTTTTTTTTTTTTTT-3'.   cDNAs
                     were cloned into the MluI/SalI sites of a modified
                     pBluescribe  vector using commercial linkers (NEB).
                     Average insert size:  1.2 kb."
BASE COUNT       26 a      3 c      6 g     17 t
ORIGIN

  Query Match             0.9%;  Score 21.8;  DB 9;  Length 52;
  Best Local Similarity   92.0%;  Pred. No. 2.3e+06;
  Matches   23;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    2117 TATATAAATATATATTTTTCAAATA 2141
           | ||||||||||||||||| |||||
Db      21 TCTATAAATATATATTTTTAAAATA 45


RESULT 57
BH792320
LOCUS       BH792320                 52 bp    DNA     linear   GSS 02-APR-2002
DEFINITION  SALK_063406.23.95.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_063406.23.95.x, DNA
            sequence.
ACCESSION   BH792320
VERSION     BH792320.1  GI:19888967
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_063406.23.95.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       15 a      18 c      10 g      9 t
ORIGIN
```

```
  Query Match             0.9%;  Score 21.8;  DB 17;  Length 52;
  Best Local Similarity   78.8%;  Pred. No. 2.3e+06;
  Matches   26;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy   1321 ATCCGATCACCTCGAAGATTTATTGGCAGCCCA 1353
          | || | | |||||||||  || ||||||||||
Db     20 ACCCTAACTCCTCGAAGAAGTAATGGCAGCCCA 52


RESULT 58
AU258750
LOCUS       AU258750                 54 bp    mRNA    linear   EST 25-APR-2002
DEFINITION  AU258750 3'-directed mouse cDNA library Mus musculus cDNA clone
            BED0013612 3', mRNA sequence.
ACCESSION   AU258750
VERSION     AU258750.1  GI:20324614
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Kato,K. and Matoba,R.
  TITLE     Generation of expressed sequence tags from mouse brain
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kikuya Kato
            Graduate School of Biological Sciences
            Nara Institute of Science and Technology
            8916-5 Takayama, Ikoma, Nara 630-0101, Japan
            Tel: 81-743-72-5581
            Fax: 81-743-72-5589
            Email: kkato@bs.aist-nara.ac.jp,
            URL:http://love2.aist-nara.ac.jp/BED/index.html.
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="BED0013612"
                     /clone_lib="3'-directed mouse cDNA library"
                     /tissue_type="brain"
                     /note="Vector: pGEM-T-easy"
BASE COUNT       18 a      7 c      7 g     21 t      1 others
ORIGIN

  Query Match             0.9%;  Score 21.8;  DB 9;  Length 54;
  Best Local Similarity   70.7%;  Pred. No. 2.2e+06;
  Matches   29;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   2099 CAACCTGAATCTTTTTTTTATATAAATATATATTTTTCAAA 2139
          | ||||| || ||||| ||  ||||||| |  |||   |||
Db     14 CGACCTGCATTTTTTTGTTTAATAAATAAAGTTTTGATAAA 54


RESULT 59
AI888083/c
LOCUS       AI888083                 55 bp    mRNA    linear   EST 01-SEP-1999
DEFINITION  wm29g07.x1 NCI_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:2437404 3',
            mRNA sequence.
ACCESSION   AI888083
VERSION     AI888083.1  GI:5593247
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2437404"
                     /clone_lib="NCI_CGAP_Ut4"
                     /tissue_type="serous papillary carcinoma, high grade, 2
                     pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.48 kb. Life Technologies catalog #:
                     11542-016"
BASE COUNT       20 a      2 c      9 g     24 t
ORIGIN

  Query Match             0.9%;  Score 21.8;  DB 9;  Length 55;
  Best Local Similarity   70.7%;  Pred. No. 2.2e+06;
  Matches   29;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   2101 ACCTGAATCTTTTTTTTATATAAATATATATTTTTCAAATA 2141
          |||||||| ||  |  || | ||||| | || | | |||||
Db     45 ACCTGAATGTTCCTACTACACAAATAAACATATATTAAATA 5


RESULT 60
AZ331030
LOCUS       AZ331030                 55 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0056G09R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0056G09 R, DNA sequence.
ACCESSION   AZ331030
VERSION     AZ331030.1  GI:10393153
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0056  row: G  column: 09
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 55.
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0056G09"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
```

/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 29 a 7 c 7 g 12 t
ORIGIN

Query Match 0.9%; Score 21.8; DB 17; Length 55;
Best Local Similarity 70.7%; Pred. No. 2.2e+06;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Qy   1798 AAATAGTATTGCTGAACTCTTAGGCACATCAATTAATTGAT 1838
          ||||| |||||   ||| ||||   | ||||| |||  |||
Db      2 AAATAATATTGAAAAACCCTTACCGAAATCAACTAAAAGAT 42
```

RESULT 61
BF791590
LOCUS BF791590 58 bp mRNA linear EST 12-JAN-2001
DEFINITION 602251706F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4344312 5', mRNA sequence.
ACCESSION BF791590
VERSION BF791590.1 GI:12096644
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 58)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM9962 row: k column: 01
High quality sequence stop: 58.
FEATURES Location/Qualifiers
source 1. .58
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4344312"
/clone_lib="NIH_MGC_84"
/tissue_type="adrenal cortex carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.229 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 49 a 1 c 2 g 6 t
ORIGIN

Query Match 0.9%; Score 21.8; DB 12; Length 58;
Best Local Similarity 65.3%; Pred. No. 2.2e+06;
Matches 32; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```
Qy   1755 ATTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATAG 1803
          || |||||||| ||  ||       |||||||||   |||  |||| ||
Db     10 ATGTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG 58
```

RESULT 62
BI846606/c
LOCUS BI846606 58 bp mRNA linear EST 04-OCT-2001
DEFINITION fq80g06.x1 Zebrafish neuronal Danio rerio cDNA clone 4886602 3', mRNA sequence.
ACCESSION BI846606
VERSION BI846606.1 GI:15959129
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes ; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 58)
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE WashU Zebrafish EST Project 1998
JOURNAL Unpublished (1998)
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
cDNA Library constructed by S. Lin DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: T7 from Gibco.
FEATURES Location/Qualifiers
source 1. .58
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="4886602"
/clone_lib="Zebrafish neuronal"
/sex="mixed"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: pBluescript (modified); Site_1: DraIII(Y); Site_2: DRAIII(X); Library is cloned directionally between the DraIII(X) and DraIII(Y) sites and has been amplified. Library constructed by S. Lin. "
BASE COUNT 8 a 4 c 2 g 44 t
ORIGIN

Query Match 0.9%; Score 21.8; DB 13; Length 58;
Best Local Similarity 61.4%; Pred. No. 2.2e+06;
Matches 35; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy   1744 AATAAACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAA 1800
          ||||||      ||  ||  ||| ||  ||||    |||||||||   |||  ||||
Db     57 AATAAAATAATGTTAAAAGCAATAAAATAATGAAAAAAAAAAAAAAAAAAAAGAAAA 1
```

RESULT 63
AA726911/c
LOCUS AA726911 59 bp mRNA linear EST 02-JAN-1998
DEFINITION vu37b05.r1 Stratagene mouse Tcell 937311 Mus musculus cDNA clone

```
            IMAGE:1193553 5', mRNA sequence.
ACCESSION   AA726911
VERSION     AA726911.1  GI:2744618
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:640649
            Seq primer: -28m13 rev1 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1193553"
                     /clone_lib="Stratagene mouse Tcell 937311"
                     /tissue_type="Tcell"
                     /dev_stage="M30 CD4+ cells"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: blood; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. M30 CD4+ cells. Average insert size: 1.0 kb;
                     Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       31 a      4 c      4 g     20 t
ORIGIN

  Query Match             0.9%;  Score 21.8;  DB 9;  Length 59;
  Best Local Similarity  61.4%;  Pred. No. 2.2e+06;
  Matches   35;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy   2112 TTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCAGCTCATTATGAAAA 2168
          |||||| | |   | | ||      ||||  || |||  ||||   ||| |||||||
Db     59 TTTTTTTTTTTTTTTTTTAAGAGAAAAATTCATATTTATTTCAAAGCATCATGAAAA 3


RESULT 64
AA547914
LOCUS       AA547914                 60 bp    mRNA    linear   EST 22-SEP-1997
DEFINITION  MB3D6V2G01T3 Brugia malayi day 6 post-infection third stage larvae
            SAW96MLW-BmL3d6 Brugia malayi cDNA clone 3D6V2G01 5', mRNA
            sequence.
ACCESSION   AA547914
VERSION     AA547914.1  GI:2315115
KEYWORDS    EST.
SOURCE      Brugia malayi.
  ORGANISM  Brugia malayi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Brugia.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Blaxter,M.L., Waterfall,M., Daub,J., Lizotte-Waniewski,M., Baron,L.
            and Jones,S.J.
  TITLE     Genes expressed in day six post-infection, third stage larvae of
            Brugia malayi
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Blaxter ML
            Institute of Cell, Animal and Population Biology
            University of Edinburgh
            Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
            3JT, UK.
            Tel: +44 131 650 6760
            Fax: +44 131 670 5450
            Email: mark.blaxter@ed.ac.uk
            The ABI trace of this sequence can be viewed at
            http://www.sanger.ac.uk/brugia/3D6/MB3D6V2G01T3.html
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Brugia malayi"
                     /strain="TRS Labs"
                     /db_xref="taxon:6279"
                     /clone="3D6V2G01"
                     /clone_lib="Brugia malayi day 6 post-infection third stage
                     larvae SAW96MLW-BmL3d6"
                     /sex="mixed"
                     /dev_stage="third stage larvae, six days after infection"
                     /lab_host="E. coli XL1-Blue"
                     /note="Vector: lambdaZapII (UniZap XR); Site_1: Eco R I
                     (5' end); Site_2: Xho I (3' end); Brugia malayi is a
                     lymphatic filarial nematode parasite of humans. mRNA was
                     prepared from third stage larvae of Brugia malayi isolated
                     from the peritoneal cavity of jirds six days after
                     infection. The mRNA was converted to double stranded cDNA
                     using reverse transcriptase and oligo(dT) followed by
                     RNase H and DNApol I. The library had 2 x 10E5 independent
                     recombinants and average insert size was 900 base pairs.
                     The library was constructed by Michelle Lizotte-Waniewski.
                     The library is available from The Filarial Genome Project
                     Resource Center: contact Dr. S.A. Williams, Clark Science
                     Center, Smith College, Northampton, MA 01063 USA phone +1
                     413 585 3826 fax +1 413 585 3786 email genome@smith.edu."
BASE COUNT       14 a      0 c      2 g     44 t
ORIGIN

  Query Match             0.9%;  Score 21.8;  DB 9;  Length 60;
  Best Local Similarity  70.7%;  Pred. No. 2.2e+06;
  Matches   29;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   2111 TTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGAT 2151
          | |||||| | | ||||||||||||      | ||||| ||
Db     12 TGTTTTTAAAAATATATATATTTTTTTTATGTTTTTTTAT 52


RESULT 65
BE896253
LOCUS       BE896253                 45 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601438976F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3923778 5',
            mRNA sequence.
ACCESSION   BE896253
VERSION     BE896253.1  GI:10360469
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9760  row: h  column: 19
```

```
          High quality sequence stop: 45.
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3923778"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
BASE COUNT       28 a      5 c      2 g     10 t
ORIGIN

  Query Match             0.9%;  Score 21.6;  DB 12;  Length 45;
  Best Local Similarity   68.2%;  Pred. No. 2.5e+06;
  Matches   30;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1756 TTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAA 1799
           || || |||| |  | |||| |  ||||| ||||  |||  |||
Db      2 TTGTACAAAAACTCTTAATGATTAAAAAACAAACAAAAAAAAAA 45


RESULT 66
AU107554
LOCUS       AU107554                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU107554 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            NBLL0035, mRNA sequence.
ACCESSION   AU107554
VERSION     AU107554.1  GI:13557075
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NBLL0035"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT       15 a     12 c     18 g      5 t
ORIGIN

  Query Match             0.9%;  Score 21.6;  DB 9;  Length 50;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+06;
  Matches   27;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy   1494 TGGGGAAGCATCGGCACCACTTCCAATACGACAGCC 1529
           |||||||| | | ||||||  ||||| | | | |||
Db      5 TGGGGAAGGAGCAGCACCAAATCCAAGATGGCGGCC 40


RESULT 67
AU009949/c
LOCUS       AU009949                 53 bp    mRNA    linear   EST 31-JUL-1998
DEFINITION  AU009949 Schizosaccharomyces pombe late log phase cDNA
            Schizosaccharomyces pombe cDNA clone spc00709, mRNA sequence.
ACCESSION   AU009949
VERSION     AU009949.1  GI:3346629
KEYWORDS    EST.
SOURCE      fission yeast.
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 53)
  AUTHORS   Morimyo,M. and Mita,K.
  TITLE     Identification of expressed sequence tags of Schizosaccharomyces
            pombe
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mitsuoki Morimyo
            Genome Research Group
            National Institute of Radiological Sciences
            9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
            Email: morimyo@nirs.go.jp.
FEATURES             Location/Qualifiers
     source          1. .53
                     /organism="Schizosaccharomyces pombe"
                     /strain="972"
                     /db_xref="taxon:4896"
                     /clone="spc00709"
                     /clone_lib="Schizosaccharomyces pombe late log phase cDNA"
                     /sex="h minus"
                     /note="Vector: M13mp19; The cDNA library of
                     Schizosaccharomyces pombe was prepared by cloning cDNA
                     into the SmaI site of M13mp19 DNA and the direction of DNA
                     sequences was not always from 5' to 3'. The cDNA data of
                     Schizosaccharomyces pombe are available for searching on
                     the World Wide Web.  (URL, http://www.nirs.go.jp)"
BASE COUNT       27 a      5 c      3 g     18 t
ORIGIN

  Query Match             0.9%;  Score 21.6;  DB 9;  Length 53;
  Best Local Similarity   63.5%;  Pred. No. 2.5e+06;
  Matches   33;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy   2110 TTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCAGCTCATT 2161
           |||||||| |  ||||||  |   | || ||| |||| |   | || |  ||
Db     53 TTTTTTTTTTGCAAATATTAAAAATACACATATATTTATTTATAAGGTGGTT 2


RESULT 68
BQ565734
LOCUS       BQ565734                 54 bp    mRNA    linear   EST 19-JUN-2002
DEFINITION  gi44e07.y1 Mouse Organ of Corti cDNA pBluescript Mus musculus cDNA
            clone gi44e07 5', mRNA sequence.
ACCESSION   BQ565734
VERSION     BQ565734.1  GI:21469051
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Kachar,B.
  TITLE     EST analysis of gene expression in the mouse Organ of Corti at the
            onset of hearing
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kachar,B.
            Structural Cell Biology
            National Institute of Deafness and other Communication Disorders
            50/4249 South Drive, NIH, Bethesda, MD 20892-8027, USA
            Tel: 301-402-1599
            Fax: 301-402-1765
            Email: kacharb@nidcd.nih.gov
```

```
          Plate: 44  row: e  column: 07
          Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /db_xref="taxon:10090"
                     /clone="gi44e07"
                     /clone_lib="Mouse Organ of Corti cDNA pBluescript"
                     /sex="male and female"
                     /dev_stage="Post natal day 5 to 13"
                     /note="Organ: Organ of Corti; Vector: pBluescript; The
                     organ of Corti (OC) was fine dissected from a total of 386
                     OC as follows: 102 samples from post-natal (P) day 5; 72
                     from P6; 60 from P7; 46 from P8; 18 from P9; 20 from P10;
                     14 from P12 and 24 from P13. After killing animals by
                     cervical dislocation followed by decapitation, the bulla
                     was removed and opened in Leibowitz medium. The bony
                     capsule of the cochlea was chipped away, stria vascularis
                     and spiral ligament were removed and the sensory
                     epithelium was carefully dissected out of the modiolus.
                     Total RNA was extracted using the micro Fasttrack kit
                     (catalog # K1593-02; Invitrogen, Carlsbad, CA), according
                     to manufacturer's instructions. Reverse transcription and
                     library construction were carried out with the Uni-Zap XR
                     vector kit (catalog # 237211, Stratagene) and Uni-Zap XR
                     Gigapack III Gold Cloning kit (catalog # 237612), both
                     from Stratagene (La Jolla, CA, USA), according to
                     manufacturer's instructions. Briefly: 1.5 ug mRNA was
                     reverse transcribed using a hybrid oligo(dT) linker-primer
                     that contains an Xho I site. First strand synthesis was
                     primed with the linker- primer and transcribed using
                     Moloney murine leukemia virus reverse transcriptase
                     (MMLV-RT) and 5-methyl dCTP. The second strand was
                     synthesized with DNA polymerase and RNase H. Complementary
                     DNA was blunt ended with Pfu DNA polymerase, ligated with
                     EcoR I adapters in the presence of ligase and digested
                     with Xho I. The cDNA was sequentially size fractionated
                     over Pharmacia Size Sep400 (Pharmacia, Uppsala, Sweden)
                     and Clontech Chroma Spin-1000 (Clontech, Palo Alto, CA)
                     columns to enrich for cDNAs greater than 400bp and 1000 bp
                     , respectively. The cDNA was then directionally ligated to
                     the Uni-ZAP XR vector, which had been predigested with
                     EcoR I and Xho I. The phagemid was packaged with Gigapak
                     III Gold and, upon titration on XL1 Blue MRF' cells, the
                     yield of the phage library was estimated to be 11,100,000
                     recombinants. Stratagene's ExAssist Interference
                     resistance helper phage (catalogue # 211203) was adopted
                     to rescue plasmid DNA from the phages. Upon plating of the
                     rescued library, individual cDNA clones were selected and
                     grown in 96-well, 2 ml growth plate. Plasmid DNA was
                     purified from 200 ul of saturated culture with the
                     Concert96(TM) plasmid purification kit (Invitrogen,
                     Carlsbad, CA) as instructed by the manufacturer. ESTs from
                     the 5' end of the cDNA clones were generated with the
                     universal M13 reverse primer (CAGGAAACAGCTATGACC) and 25%
                     strength BigDye terminator sequencing chemistry (Applied
                     Biosystems, Foster City, CA). Sequencing reactions were
                     performed on MJ Tetrad thermal cyclers (MJ Research,
                     Waltham, MA), and analyzed on 3700 automated capillary
                     sequencers using POP5 polymer (Applied Biosystems, Foster
                     City, CA).    The frequency distribution of the library is
                     as follows: 72% of genes have 1 copy; 14.3% 2; 12% 3-10;
                     1.4% 11-50 and 0.1% 51-150. As to gene function, 45% of
                     genes are present in GenBank and have know function; 23%
                     have hits in GenBank, but do not have assigned function;
                     12% are uncharacterized ESTs and 20% are unidentified."
BASE COUNT       16 a      9 c      26 g       3 t
ORIGIN

  Query Match             0.9%;  Score 21.6;  DB 14;  Length 54;
  Best Local Similarity   63.5%;  Pred. No. 2.5e+06;
  Matches   33;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy    117 AGACCTGGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGG 168
          ||||  |||     |||  ||||  || || ||  || |||   ||||||| |
Db      3 AGACGGGGAGGCACCACTGGACAAGGCCTGTGCAGAGGAGGAAGAGGAGGAG 54


RESULT 69
BH846951
LOCUS       BH846951                 54 bp    DNA     linear   GSS 13-JUN-2002
DEFINITION  SALK_012126.50.20.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_012126.50.20.x, DNA
            sequence.
ACCESSION   BH846951
VERSION     BH846951.1  GI:21417822
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_012126.50.20.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       19 a      2 c      5 g      28 t
ORIGIN

  Query Match             0.9%;  Score 21.6;  DB 17;  Length 54;
  Best Local Similarity   68.2%;  Pred. No. 2.5e+06;
  Matches   30;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   2111 TTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCA 2154
           ||||| ||  | |||   ||||||  ||||| || ||  |||
Db      7 TTTTTATAGTTTAATGACTATTTTCTAAATAAATGTTATATTTA 50


RESULT 70
AL786691
LOCUS       AL786691                 55 bp    mRNA    linear   EST 27-JUN-2002
DEFINITION  AL786691 XGC-neurula Silurana tropicalis cDNA clone TNeu096m20 5',
            mRNA sequence.
ACCESSION   AL786691
VERSION     AL786691.1  GI:21572395
KEYWORDS    EST.
SOURCE      western clawed frog.
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Taylor,R., Ashurst,J.L., Croning,M.D.R., Zorn,A.M. and Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2002
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Taylor R
            Sanger Centre
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TNeu096m20.p1cSP6
            Sequencing primer: P1CSP6
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
FEATURES              Location/Qualifiers
     source           1. .55
                      /organism="Silurana tropicalis"
                      /db_xref="taxon:8364"
                      /clone="TNeu096m20"
                      /clone_lib="XGC-neurula"
                      /dev_stage="neurula"
                      /lab_host="Escherichia coli DH10B"
                      /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                      was oligo dT primed from 5ug of poly A+ RNA from neurula.
                      EcoRI-NotI cut cDNA was then ligated into pCS107 with
                      EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT       43 a      4 c      5 g      3 t
ORIGIN

  Query Match             0.9%;  Score 21.6;  DB 9;  Length 55;
  Best Local Similarity   63.5%;  Pred. No. 2.5e+06;
  Matches   33;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy     903 ATTCACTGGGGAGAATAGAAAGAAAACAATTGACAAAATCCTCAAATGTAAA 954
           |||| | ||||| ||.| ||| |||| ||   | ||||     |||   |||
Db       3 ATTCCCCGGGGAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAA 54


RESULT 71
AU264938/c
LOCUS       AU264938                 56 bp    mRNA    linear   EST 10-MAY-2002
DEFINITION  AU264938 VS Dictyostelium discoideum cDNA clone VSF251 5', mRNA
            sequence.
ACCESSION   AU264938
VERSION     AU264938.1  GI:20523736
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   Urushihara,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
            Takeuchi,I., Kohara,Y. and Tanaka,Y.
  TITLE     Population analysis of cDNAs from unicellular and multicellular
            stages of Dictyostelium discoideum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.
FEATURES              Location/Qualifiers
     source           1. .56
                      /organism="Dictyostelium discoideum"
                      /strain="AX4"
                      /db_xref="taxon:44689"
                      /clone="VSF251"
                      /clone_lib="VS"
                      /sex="mat A"
                      /dev_stage="vegetative"
BASE COUNT       36 a      2 c      0 g     16 t      2 others
ORIGIN

  Query Match             0.9%;  Score 21.6;  DB 9;  Length 56;
  Best Local Similarity   61.1%;  Pred. No. 2.5e+06;
  Matches   33;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy   2110 TTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCAGCTCATTAT 2163
          |||||||| || ||| ||   ||||| |  |  |  |||  || |  | | | |
Db     56 TTTTTTTTTTAAAAAAATNGTTTTTTTATTTTAAAATTTNGTTTATTTTAATTT 3


RESULT 72
BI855869/c
LOCUS       BI855869                 56 bp    mRNA    linear   EST 10-OCT-2001
DEFINITION  603383029F2 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5391846 5',
            mRNA sequence.
ACCESSION   BI855869
VERSION     BI855869.1  GI:15996616
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11998  row: j  column: 07
            High quality sequence stop: 56.
FEATURES              Location/Qualifiers
     source           1. .56
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:5391846"
                      /clone_lib="NIH_MGC_87"
                      /tissue_type="mammary adenocarcinoma, cell line"
                      /lab_host="DH10B (phage-resistant)"
                      /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                      Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                      Average insert size 1.383 kb. Library enriched for
                      full-length clones and constructed by Life Technologies.
                      Note: this is a NIH_MGC Library."
BASE COUNT       31 a      1 c      5 g     19 t
ORIGIN

  Query Match             0.9%;  Score 21.6;  DB 13;  Length 56;
  Best Local Similarity   63.5%;  Pred. No. 2.5e+06;
  Matches   33;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCAGCTCA 2159
          | |||||||| ||  || || |||  ||  || |  ||| |  | ||  |||
Db     54 TTTTTTTTTTTTAATAAAATGTATCCTTATAAAAATTTTATTCTACATTTCA 3


RESULT 73
BI705223/c
LOCUS       BI705223                 58 bp    mRNA    linear   EST 19-SEP-2001
DEFINITION  fr66e10.y1 Zebrafish SJD day 8 fin regeneration Danio rerio cDNA
            clone 4962547 5', mRNA sequence.
ACCESSION   BI705223
VERSION     BI705223.1  GI:15680918
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes ; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 58)
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE WashU Zebrafish EST Project 1998
JOURNAL Unpublished (1998)
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
cDNA Library construction by: Joe Barnes and Steve Johnson. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Research Genetics web address:
http://www.researchgenetics.com/
Putative full length read
The vector to vector length is 59
Seq primer: T3 ET from Amersham.
FEATURES Location/Qualifiers
source 1. .58
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="4962547"
/clone_lib="Zebrafish SJD day 8 fin regeneration"
/sex="male"
/tissue_type="fin, 8-day regeneration"
/lab_host="DH10B"
/note="Vector: pAMP1; Site_1: EcoRI; Site_2: NotI; First strand cDNA synthesis was primed using oligo-dT on magnetic beads with an additional primer 5'-gcggccgctaatacgactcacta-taggg-3'. Second strand synthesis was a 3-cycle PCR using the primers 5'-ggccgctaatacgactcactatag-3' and 5'-aagcagtggtaacaacgcagagtacttt-ttttttttttttvn-3'. cDNA was subsequently amplified in a 7-cycle PCR with the following primers: 5'-ggccgctaatacgactcactatag-3' and 5'-aagcagtggt-aacaacgcag. Deoxy-UMP adaptors were added in a third PCR (5 cycles) and the primers 5'-caucaucaucauggccgctaatacgactcactataggg-3' and 5'-cuacuacuacuaaagcagtggtaacaacgcagagtac-3'. Ends were treated with uracil DNA glycosylase and product with 3' overhangs was annealed to complementary ends of pAMP1. Insert can be excised using EcoRI and NotI. Library constructed by Joe Barnes and Steve Johnson (Washington University)."
BASE COUNT 41 a 0 c 5 g 12 t
ORIGIN

Query Match 0.9%; Score 21.6; DB 13; Length 58;
Best Local Similarity 63.5%; Pred. No. 2.5e+06;
Matches 33; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 2110 TTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCAGCTCATT 2161
||||||||| | | | ||||||| ||| |||| | ||| |||
Db 58 TTTTTTTTTTTTTTTTTTATATTTAAACAATTTATTTATTATTTTTTCCATT 7

RESULT 74
BI709260/c
LOCUS BI709260 58 bp mRNA linear EST 19-SEP-2001
DEFINITION fs62b03.y1 Zebrafish SJD day 8 fin regeneration Danio rerio cDNA clone 5072212 5', mRNA sequence.
ACCESSION BI709260
VERSION BI709260.1 GI:15684955
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes ; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 58)
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE WashU Zebrafish EST Project 1998
JOURNAL Unpublished (1998)
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
cDNA Library construction by: Joe Barnes and Steve Johnson. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Research Genetics web address:
http://www.researchgenetics.com/
Putative full length read
The vector to vector length is 59
Seq primer: T3 ET from Amersham.
FEATURES Location/Qualifiers
source 1. .58
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="5072212"
/clone_lib="Zebrafish SJD day 8 fin regeneration"
/sex="male"
/tissue_type="fin, 8-day regeneration"
/lab_host="DH10B"
/note="Vector: pAMP1; Site_1: EcoRI; Site_2: NotI; First strand cDNA synthesis was primed using oligo-dT on magnetic beads with an additional primer 5'-gcggccgctaatacgactcacta-taggg-3'. Second strand synthesis was a 3-cycle PCR using the primers 5'-ggccgctaatacgactcactatag-3' and 5'-aagcagtggtaacaacgcagagtacttt-ttttttttttttvn-3'. cDNA was subsequently amplified in a 7-cycle PCR with the following primers: 5'-ggccgctaatacgactcactatag-3' and 5'-aagcagtggt-aacaacgcag. Deoxy-UMP adaptors were added in a third PCR (5 cycles) and the primers 5'-caucaucaucauggccgctaatacgactcactataggg-3' and 5'-cuacuacuacuaaagcagtggtaacaacgcagagtac-3'. Ends were treated with uracil DNA glycosylase and product with 3' overhangs was annealed to complementary ends of pAMP1. Insert can be excised using EcoRI and NotI. Library constructed by Joe Barnes and Steve Johnson (Washington University)."
BASE COUNT 41 a 0 c 5 g 12 t
ORIGIN

Query Match 0.9%; Score 21.6; DB 13; Length 58;
Best Local Similarity 63.5%; Pred. No. 2.5e+06;
Matches 33; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 2110 TTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCAGCTCATT 2161
||||||||| | | | ||||||| ||| |||| | ||| |||
Db 58 TTTTTTTTTTTTTTTTTTATATTTAAACAATTTATTTATTATTTTTTCCATT 7

RESULT 75
BI496827
LOCUS BI496827 59 bp mRNA linear EST 28-AUG-2001
DEFINITION df128f02.w1 Morton Fetal Cochlea Homo sapiens cDNA clone IMAGE:2538387 3', mRNA sequence.
ACCESSION BI496827
VERSION BI496827.1 GI:15336171
KEYWORDS EST.

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Robertson,N.G., Khetarpal,U., Gutierrez-Espeleta,G.A., Bieber,F.R.
            and Morton,C.C.
  TITLE     Isolation of novel and known genes from a human fetal cochlear cDNA
            library using subtractive hybridization and differential screening
  JOURNAL   Genomics 23, 42-50 (1994)
  MEDLINE   95130111
COMMENT     Contact: Morton, C. C.
            Departments of Pathology and Obstetrics, Gynecology and
            Reproductive Biology
            Brigham and Women's Hospital
            75 Francis Street, Harvard Medical School, Boston, MA 02115, USA
            Tel: 617 732 7980
            Fax: 617 738 6996
            Email: ccmorton@bics.bwh.harvard.edu
            DNA sequencing and analyses were performed by National Institutes
            of Health Intramural Sequencing Center (NISC; see
            http://www.nisc.nih.gov).
             This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Plate: LLAM6322  row: L  column: 4
            Seq primer: T7 primer.
FEATURES              Location/Qualifiers
     source           1. .59
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2538387"
                      /clone_lib="Morton Fetal Cochlea"
                      /tissue_type="cochlea"
                      /dev_stage="16-22 week fetus"
                      /lab_host="SOLR cells (kanamycin resistant)"
                      /note="Organ: ear; Vector: pBluescript SK-; Site_1: EcoRI;
                      Site_2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned
                      unidirectionally. Primer: Oligo dT. Fetal cochlea, normal.
                      37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP
                      XR Vector. Library constructed by N. Robertson, C. Morton.
                      ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                      sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT        9 a     11 c      8 g     31 t
ORIGIN

  Query Match             0.9%;  Score 21.6;  DB 13;  Length 59;
  Best Local Similarity   85.7%;  Pred. No. 2.5e+06;
  Matches   24;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy    2108 TCTTTTTTTTATATAAATATATATTTTT 2135
           | |||||||||||||| |||| ||||| |
Db      24 TTTTTTTTTTATATATATATGTATTTAT 51


RESULT 76
BG348123/c
LOCUS       BG348123                 60 bp    mRNA    linear   EST 28-FEB-2001
DEFINITION  de78g05.y1 Kirschner embryo St10 14 Xenopus laevis cDNA clone
            IMAGE:3518001 5', mRNA sequence.
ACCESSION   BG348123
VERSION     BG348123.1  GI:13168549
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.
            , Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
            ,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     WashU Xenopus EST project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Sandy Clifton, Ph.D.
            WashU Xenopus EST project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library constructed by M. Kirschner (Harvard Medical School). DNA
            Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: Xenopus clones from this library are available
            through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
            Seq primer: -40RP from Gibco.
FEATURES              Location/Qualifiers
     source           1. .60
                      /organism="Xenopus laevis"
                      /db_xref="taxon:8355"
                      /clone="IMAGE:3518001"
                      /clone_lib="Kirschner embryo St10 14"
                      /tissue_type="pooled embryos (stage 10-14)"
                      /lab_host="DH10B (phage-resistant)"
                      /note="Vector: pCS2+; Site_1: NotI; Site_2: SalI;
                      Size-selected for average insert size 1.2 kb. Library was
                      constructed and donated by M. Kirschner (Harvard Medical
                      School)."
BASE COUNT       32 a      3 c      5 g     20 t
ORIGIN

  Query Match             0.9%;  Score 21.6;  DB 12;  Length 60;
  Best Local Similarity   60.0%;  Pred. No. 2.5e+06;
  Matches   36;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy    2109 CTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCAGCTCATTATGAAAA 2168
           |||||||||  | |   | |  || | | ||  |||||| |  |  | | || ||  || |
Db      60 CTTTTTTTTTTTTTTTTTTTTAATGTATAAATCAGATTTATATTATAACACAATAGCAATA 1


RESULT 77
C25778/c
LOCUS       C25778                   60 bp    mRNA    linear   EST 29-JUL-1997
DEFINITION  C25778 Dictyostelium discoideum FC (H.Urushihara) Dictyostelium
            discoideum cDNA clone FC-BC03, mRNA sequence.
ACCESSION   C25778
VERSION     C25778.1  GI:2282564
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Suzuki,K., Shimizu,H. and Urushiihara,H.
  TITLE     Sexual cDNA in D. discoideum(970724)
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.
FEATURES              Location/Qualifiers
     source           1..60
                      /organism="Dictyostelium discoideum"
                      /strain="KAX3"
                      /db_xref="taxon:44689"
                      /clone="FC-BC03"
                      /clone_lib="Dictyostelium discoideum FC (H.Urushihara)"
BASE COUNT       46 a      1 c      2 g     11 t
ORIGIN

  Query Match             0.9%;  Score 21.6;  DB 14;  Length 60;
  Best Local Similarity   68.2%;  Pred. No. 2.5e+06;
  Matches   30;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;
```

```
Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGAT 2151
          |||||||||  |  | ||| ||| |||||   ||  | | ||||
Db     50 TATTTTTTTTTTTTTAATTTATTTTTTTTTTATCTAATGTTTAT 7


RESULT 78
N81844/c
LOCUS       N81844                    60 bp    mRNA    linear   EST 22-MAY-2000
DEFINITION  TgESTzy48g03.r1 TgRH Tachyzoite cDNA Toxoplasma gondii cDNA clone
            tgzy48g03.r1 5', mRNA sequence.
ACCESSION   N81844
VERSION     N81844.1  GI:1257597
KEYWORDS    EST.
SOURCE      Toxoplasma gondii.
  ORGANISM  Toxoplasma gondii
            Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
            Sarcocystidae; Toxoplasma.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Hehl,A., Manger,I., Marra,M., Sibley,L.D., Ajioka,J.A., Aslett,M.A.
            , Dietrich,N., Dubuque,T., Hillier,L., Kucaba,T., Wan,K.L.,
            Waterston,R.H. and Boothroyd,J.
  TITLE     WashU-Merck-Stanford-NIH Toxoplasma EST project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: toxo@watson.wustl.edu
            David Sibley at toxoest@borcim.wustl.edu for further information
            relating to organism, clone or library availability.
            Seq primer: T3
            High quality sequence stop: 50.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Toxoplasma gondii"
                     /strain="RH"
                     /db_xref="taxon:5811"
                     /clone="tgzy48g03.r1"
                     /clone_lib="TgRH Tachyzoite cDNA"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: Lambda ZAP; Site_1: EcoRI; Site_2: XhoI;
                     Toxoplasma RH strain tachyzoites were grown in human
                     foreskin fibroblast cultures in vitro. The library was
                     constructed by K.L. Wan, Cambridge University.  cDNAs were
                     synthesized from polyA RNAs by oligo d(T) priming and
                     directionally cloned into the EcoRI to XhoI sites of the
                     Lambda ZapII vector using the ZAP-cDNA synthesis kit
                     (Stratagene). WARNING: the library contains a small
                     percentage of cDNAs derived from the human host cells."
BASE COUNT       10 a      0 c     11 g     39 t
ORIGIN

  Query Match             0.9%;  Score 21.6;  DB 14;  Length 60;
  Best Local Similarity   60.0%;  Pred. No. 2.5e+06;
  Matches   36;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy    916 AATAGAAAGAAAACAATTGACAAAATCCTCAAATGTAAACTCAATTTGCCTCCCTACCTC 975
          || | ||| |||| ||   | ||||     |||   ||||||  | |  ||| ||  |||
Db     60 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTCTCTCTCTCTCTCTCTC 1


RESULT 79
AZ368409/c
LOCUS       AZ368409                  41 bp    DNA     linear   GSS 02-OCT-2000
DEFINITION  1M0118A04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0118A04 R, DNA sequence.
ACCESSION   AZ368409
VERSION     AZ368409.1  GI:10482109
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 41)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0118  row: A  column: 04
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 41.
FEATURES             Location/Qualifiers
     source          1. .41
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0118A04"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       27 a      3 c      1 g     10 t
ORIGIN

  Query Match             0.9%;  Score 21.4;  DB 17;  Length 41;
  Best Local Similarity   71.8%;  Pred. No. 2.8e+06;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTT 2146
          | |||||||| | |||| || ||||| |  | || || |
Db     40 TTTTTTTTTTCTTTAAAGATTTATTTATTTAGTATATGT 2


RESULT 80
AI431480
LOCUS       AI431480                  46 bp    mRNA    linear   EST 09-MAR-1999
DEFINITION  th40f03.x1 NCI_CGAP_Lym12 Homo sapiens cDNA clone IMAGE:2120765 3'
            similar to TR:Q26195 Q26195 PVA1 GENE. ;, mRNA sequence.
ACCESSION   AI431480
VERSION     AI431480.1  GI:4303480
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Life Technologies catalog #: 11547-015
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .46
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2120765"
                      /clone_lib="NCI_CGAP_Lym12"
                      /tissue_type="lymphoma, follicular mixed small and large
                      cell"
                      /lab_host="DH10B"
                      /note="Organ: lymph node; Vector: pCMV-SPORT6; Site_1:
                      SalI; Site_2: NotI; Cloned unidirectionally.  Primer:
                      Oligo dT.  Average insert size 1.25 kb. Life Technologies
                      catalog #: 11547-015"
BASE COUNT       26 a      1 c     16 g      3 t
ORIGIN

  Query Match             0.9%;  Score 21.4;  DB 9;  Length 46;
  Best Local Similarity   80.6%;  Pred. No. 2.8e+06;
  Matches   25;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy   2041 AGGCAAAATGGGAAGGCAAAGACAAAAGAAA 2071
           ||||| || ||||||| ||||| |||  |||
Db      2 AGGCATAAAGGGAAGGTAAAGAAAAATAAAA 32


RESULT 81
BG222966
LOCUS       BG222966                 49 bp    mRNA    linear   EST 07-FEB-2001
DEFINITION  nah40b02.x1 NCI_CGAP_HN21 Homo sapiens cDNA clone IMAGE:4233530 3',
            mRNA sequence.
ACCESSION   BG222966
VERSION     BG222966.1  GI:12708487
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   NCI/NIDR-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Dental Research,
            Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco.
FEATURES              Location/Qualifiers
     source           1. .49
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:4233530"
                      /clone_lib="NCI_CGAP_HN21"
                      /tissue_type="nasopharyngeal carcinoma"
                      /lab_host="DH10B"
                      /note="Organ: head/neck; Vector: pAMP1; mRNA made from
                      head/neck tumor, cDNA made by oligo-dT  priming.
                      Directionally cloned into UDG sites. Size-selected on
                      agarose gel, average insert size 300 bp. Primary library.
                      cDNA Library Preparation: David B. Krizman, Ph.D."
BASE COUNT       11 a      11 c      9 g      18 t
ORIGIN

  Query Match             0.9%;  Score 21.4;  DB 12;  Length 49;
  Best Local Similarity   66.0%;  Pred. No. 2.8e+06;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    203 TTGGACCATATGAACTTGGCATGGAACATTGTGAGAAATTTGAAATC 249
          ||   || ||||| | || |   | |||| || ||||  ||||| ||
Db      1 TTTTTCCTTATGAGCATGTCCCCGTACATCGTTAGAAGCTTGAATTC 47


RESULT 82
BJ064587/c
LOCUS       BJ064587                 51 bp    mRNA    linear   EST 10-DEC-2001
DEFINITION  BJ064587 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL080b20 5', mRNA sequence.
ACCESSION   BJ064587
VERSION     BJ064587.1  GI:17471537
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            ,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES              Location/Qualifiers
     source           1. .51
                      /organism="Xenopus laevis"
                      /db_xref="taxon:8355"
                      /clone="XL080b20"
                      /clone_lib="NIBB Mochii normalized Xenopus tailbud
                      library"
                      /tissue_type="whole embryo"
                      /dev_stage="stage 25"
                      /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
                      were oligo-dT primed and directionally cloned. Staging
                      according to  Nieuwkoop and Faber. Library is subtracted
                      and was constructed by N. Garrett  and A.M. Zorn,
                      (Wellcome/CRC Institute). "
BASE COUNT       27 a       0 c       5 g      19 t
ORIGIN

  Query Match             0.9%;  Score 21.4;  DB 13;  Length 51;
  Best Local Similarity   71.8%;  Pred. No. 2.8e+06;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTT 2146
           | ||||| ||  | | | ||| ||||||| ||||| || |
Db     48 TTTTTTTCTTCAAAATAAATTTATTTTTTAAATATATAT 10


RESULT 83
```

```
BG271387
LOCUS       BG271387                  52 bp    mRNA    linear   EST 20-FEB-2001
DEFINITION  nai49h05.x1 NCI_CGAP_HN20 Homo sapiens cDNA clone IMAGE:4263776 3',
            mRNA sequence.
ACCESSION   BG271387
VERSION     BG271387.1  GI:12979476
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   NCI/NIDR-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Dental Research,
            Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4263776"
                     /clone_lib="NCI_CGAP_HN20"
                     /lab_host="DH10B"
                     /note="Organ: normal head/neck tissue; Vector: pAMP1; mRNA
                     made from head/neck tissue, cDNA made by oligo-dT
                     priming. Directionally cloned into UDG sites.
                     Size-selected on  agarose gel, average insert size 300 bp.
                     Primary library.  cDNA Library Preparation: David B.
                     Krizman, Ph.D."
BASE COUNT       12 a     12 c     10 g     18 t
ORIGIN

  Query Match             0.9%;  Score 21.4;  DB 12;  Length 52;
  Best Local Similarity   66.0%;  Pred. No. 2.8e+06;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    203 TTGGACCATATGAACTTGGCATGGAACATTGTGAGAAATTTGAAATC 249
          ||   || ||||| | || |    | |||| || ||||  ||||| ||
Db      1 TTTTTCCTTATGAGCATGTCCCCGTACATCGTTAGAAGCTTGAATTC 47


RESULT 84
BG231005
LOCUS       BG231005                  53 bp    mRNA    linear   EST 09-FEB-2001
DEFINITION  nai43b05.x1 NCI_CGAP_HN20 Homo sapiens cDNA clone IMAGE:4262744 3',
            mRNA sequence.
ACCESSION   BG231005
VERSION     BG231005.1  GI:12726083
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 53)
  AUTHORS   NCI/NIDR-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Dental Research,
            Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .53
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4262744"
                     /clone_lib="NCI_CGAP_HN20"
                     /lab_host="DH10B"
                     /note="Organ: normal head/neck tissue; Vector: pAMP1; mRNA
                     made from head/neck tissue, cDNA made by oligo-dT
                     priming. Directionally cloned into UDG sites.
                     Size-selected on  agarose gel, average insert size 300 bp.
                     Primary library.  cDNA Library Preparation: David B.
                     Krizman, Ph.D."
BASE COUNT       10 a     11 c      8 g     24 t
ORIGIN

  Query Match             0.9%;  Score 21.4;  DB 12;  Length 53;
  Best Local Similarity   66.0%;  Pred. No. 2.8e+06;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    203 TTGGACCATATGAACTTGGCATGGAACATTGTGAGAAATTTGAAATC 249
          ||   || ||||| | || |    | |||| || ||||  ||||| ||
Db      7 TTTTTCCTTATGAGCATGTCCCCGTACATCGTTAGAAGCTTGAATTC 53


RESULT 85
AZ366215
LOCUS       AZ366215                  55 bp    DNA     linear   GSS 02-OCT-2000
DEFINITION  1M0115A14F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0115A14 F, DNA sequence.
ACCESSION   AZ366215
VERSION     AZ366215.1  GI:10479915
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0115  row: A  column: 14
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 55.
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0115A14"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
```

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 7 a 0 c 2 g 46 t
ORIGIN

Query Match 0.9%; Score 21.4; DB 17; Length 55;
Best Local Similarity 66.0%; Pred. No. 2.8e+06;
Matches 31; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCA 2154
          | |||||||| | |    | | | ||||| ||| |  || || ||| |
Db      7 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAAGATGTTATTTATTTA 53
```

RESULT 86
N63590/c
LOCUS N63590 57 bp mRNA linear EST 01-MAR-1996
DEFINITION yz77a05.s1 Soares_multiple_sclerosis_2NbHMSP Homo sapiens cDNA clone IMAGE:289040 3', mRNA sequence.
ACCESSION N63590
VERSION N63590.1 GI:1211419
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 57)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: m13 -40 forward
High quality sequence stop: 384.
FEATURES Location/Qualifiers
source 1. .57
/organism="Homo sapiens"
/db_xref="GDB:3904880"
/db_xref="taxon:9606"
/clone="IMAGE:289040"
/clone_lib="Soares_multiple_sclerosis_2NbHMSP"
/sex="male"
/tissue_type="multiple sclerosis lesions"
/dev_stage="Age 46"
/lab_host="DH10B (ampicillin resistant)"
/note="Vector: pT7T3D (Pharmacia) with a modified polylinker V_TYPE: phagemid; Site_1: Not I; Site_2: Eco RI ; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Patima Bonaldo. RNA from 4 multiple sclerosis lesions from one patient was kindly provided by Dr. Kevin G. Becker (NINDS/NIH). "

BASE COUNT 18 a 8 c 11 g 20 t
ORIGIN

Query Match 0.9%; Score 21.4; DB 14; Length 57;
Best Local Similarity 66.0%; Pred. No. 2.8e+06;
Matches 31; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy   1764 AATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATAGTATTGCT 1810
          || ||  | | ||||||    | | ||  || ||||||||||| ||
Db     55 AAAGAACTATTTGTGCAACTCCAGATTTTCAGTAAAATAGTATTACT 9
```

RESULT 87
B03506/c
LOCUS B03506 57 bp DNA linear GSS 13-JUL-1996
DEFINITION cSRL-180D5-u cSRL flow sorted Chromosome 11 specific cosmid Homo sapiens genomic clone cSRL-180D5, DNA sequence.
ACCESSION B03506
VERSION B03506.1 GI:1412784
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 57)
AUTHORS Evans,G.A., Burbee,D., Davies,C., Hahner,L., Oliver,T., Gilbert,M., Jones,D., Ward,T., Gillilan,E., Schagemann,J., Probst,S. ., Harris,J., DeFord,J., McFarland,J., Burzinski,K., Khan,M., Kupfer,K. and Garner,H.R.
TITLE Genomic Sequence Sampled Map of Chromosome 11
JOURNAL Unpublished (1996)
COMMENT Contact: Evans GA, Shane Probst
McDermott Center for Human Growth and Development
University of Texas Southwestern Medical Center At Dallas
5323 Harry Hines Blvd, Dallas TX 75235-8591
Tel: 214-648-1600
Fax: 214-648-1666
Email: gevans@utsw.swmed.edu, shane@mcdermott.swmed.edu
Seq primer: T7
Class: cosmid ends
High quality sequence stop: 57.
FEATURES Location/Qualifiers
source 1. .57
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="cSRL-180D5"
/clone_lib="cSRL flow sorted Chromosome 11 specific cosmid"
/sex="female"
/cell_type="chimeric hamster somatic cell hybrid"
/note="Vector: sCos-1; Human Chromosome 11 specific cosmid library prepared from flow sorted human Chromosome 11 derived from Chinese Hampster Ovary (CHO) monochromosomal somatic cell hybrid, J1"
BASE COUNT 31 a 1 c 5 g 19 t 1 others
ORIGIN

Query Match 0.9%; Score 21.4; DB 17; Length 57;
Best Local Similarity 64.6%; Pred. No. 2.8e+06;
Matches 31; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```
Qy   2225 TTCTGTTTCTTCTCTTGGTGAAATAATAAAATGCAAATGAATCATTGT 2272
          || | ||| || |||   | ||||| |||||   |||   | ||| |
Db     48 TTTTTTTTTTTTTCTNATTCAAATAAGAAAATAATAATTTTTAATTCT 1
```

```
RESULT 88
AI473426/c
LOCUS       AI473426                 59 bp    mRNA    linear   EST 13-APR-1999
DEFINITION  tj13g01.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2141424 3',
            mRNA sequence.
ACCESSION   AI473426
VERSION     AI473426.1  GI:4326471
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1400   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 58.
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2141424"
                     /clone_lib="NCI_CGAP_Gas4"
                     /tissue_type="poorly differentiated adenocarcinoma with
                     signet ring cell features"
                     /lab_host="DH10B"
                     /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.69 kb. Life Technologies catalog #:
                     11549-011"
BASE COUNT        8 a      0 c      2 g     49 t
ORIGIN

  Query Match             0.9%;  Score 21.4;  DB 9;  Length 59;
  Best Local Similarity   66.0%;  Pred. No. 2.8e+06;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   1756 TTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
           ||||||||||  ||  ||       |||||||||   |||  ||||  |
Db     53 TTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 7


RESULT 89
AL769947/c
LOCUS       AL769947                 59 bp    DNA     linear   GSS 19-JUN-2002
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-095A02-012025,
            genomic survey sequence.
ACCESSION   AL769947
VERSION     AL769947.1  GI:21532149
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 59)
  AUTHORS   Rosso,M., Li,Y., Strizhov,N. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion close to or within gene At2g23880. The
            sequences are generated at the MPI for Plant Breeding Research in
            the context of the GABI-Kat project. GABI-Kat is part of the German
            Plant Genomics program designated 'GABI'. Information on line
            availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-095A02-012025"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161. The lines contain one or more T-DNA
                     insertions. The DNA fragment(s) resulting from the PCR
                     were directly sequenced to determine the genomic sequence
                     flanking the insertion. Sequences displaying significant
                     similarity to the A. thalaina nuclear genome sequence were
                     processed for submission. T-DNA derived sequences were
                     removed"
BASE COUNT       10 a      8 c      7 g     34 t
ORIGIN

  Query Match             0.9%;  Score 21.4;  DB 17;  Length 59;
  Best Local Similarity   71.8%;  Pred. No. 2.8e+06;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   2044 CAAAATGGGAAGGCAAAGACAAAAGAAACTTACCAATTG 2082
           |||  ||||||  ||||  | ||||  |||    ||  |||||
Db     57 CAATTTGGGAATCCAAAAAAAAAAAAAAAAAAACAAATTG 19


RESULT 90
AV837933/c
LOCUS       AV837933                 39 bp    mRNA    linear   EST 07-NOV-2001
DEFINITION  AV837933 Nori Satoh unpublished cDNA library, cleavage stage embryo
            Ciona intestinalis cDNA clone rcicl01m15, mRNA sequence.
ACCESSION   AV837933
VERSION     AV837933.1  GI:16782084
KEYWORDS    EST.
SOURCE      Ciona intestinalis.
  ORGANISM  Ciona intestinalis
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
  TITLE     Expressed genes in Ciona intestinalis
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Nori Satoh
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4081
            Fax: 81-75-705-1113
            Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .39
```

```
                     /organism="Ciona intestinalis"
                     /db_xref="taxon:7719"
                     /clone="rcicl01m15"
                     /clone_lib="Nori Satoh unpublished cDNA library, cleavage
                     stage embryo"
                     /tissue_type="whole animal"
                     /dev_stage="cleavage stage embryo"
                     /note="Vector: pBluescript SK"
BASE COUNT       15 a      5 c      2 g     16 t      1 others
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 10;  Length 39;
  Best Local Similarity   88.5%;  Pred. No. 3.1e+06;
  Matches   23;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   1743 AAATAAACGTGGATTTTAAAAAATCA 1768
          ||||| | ||| ||||||||||||||
Db     31 AAATACAAGTGTATTTTAAAAAATCA 6


RESULT 91
BH235137
LOCUS       BH235137                 39 bp    DNA     linear   GSS 01-JAN-2002
DEFINITION  MSAD_E07.x S Spiroplasma kunkelii genomic clone MSAD_E07.x, DNA
            sequence.
ACCESSION   BH235137
VERSION     BH235137.1  GI:18030605
KEYWORDS    GSS.
SOURCE      Spiroplasma kunkelii.
  ORGANISM  Spiroplasma kunkelii
            Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
            Spiroplasmataceae; Spiroplasma.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Hogenhout,S.A.
  TITLE     Genomic sequences from Spiroplasma kunkelii strain M2
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Hogenhout SA
            Department of Entomology
            The Ohio State University-OARDC
            120 Thorne Hall, 1680 Madison Avenue, Wooster, OH 44691, USA
            Tel: 330 263 3730
            Fax: 330 263 3686
            Email: hogenhout.1@osu.edu
            Plate: AD  row: E  column: 07
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="Spiroplasma kunkelii"
                     /strain="M2"
                     /db_xref="taxon:47834"
                     /clone="MSAD_E07.x"
                     /clone_lib="S"
BASE COUNT       14 a      0 c      1 g     24 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 17;  Length 39;
  Best Local Similarity   76.5%;  Pred. No. 3.1e+06;
  Matches   26;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   2107 ATCTTTTTTTTATATAAATATATATTTTTCAAAT 2140
          || ||||| |  |||||||| ||||||||   ||
Db      1 ATATTTTTATATTATAAATAAATATTTTTATTAT 34


RESULT 92
AA027616/c
LOCUS       AA027616                 46 bp    mRNA    linear   EST 21-JAN-1997
DEFINITION  mi08c08.r1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA
            clone IMAGE:459854 5' similar to SW:SPS1_YEAST P08458
            SPORULATION-SPECIFIC PROTEIN 1 ;, mRNA sequence.
ACCESSION   AA027616
VERSION     AA027616.1  GI:1493610
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:276742
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:459854"
                     /clone_lib="Soares mouse placenta 4NbMP13.5 14.5"
                     /sex="unknown"
                     /tissue_type="placenta"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: placenta; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       12 a      8 c     14 g     12 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 9;  Length 46;
  Best Local Similarity   69.0%;  Pred. No. 3.1e+06;
  Matches   29;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    778 CACACATTTTGTGGAACAATAGAATACATGGCCCCTGAAATC 819
          || || |||  || |||   | | |  ||||||||||| ||
Db     43 CAAACCTTTGTTGTAACCCCATATTGGATGGCCCCTGAAGTC 2


RESULT 93
BG271415
LOCUS       BG271415                 49 bp    mRNA    linear   EST 20-FEB-2001
DEFINITION  nai50d11.x1 NCI_CGAP_HN20 Homo sapiens cDNA clone IMAGE:4263597 3',
            mRNA sequence.
ACCESSION   BG271415
VERSION     BG271415.1  GI:12979533
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   NCI/NIDR-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Dental Research,
```

```
            Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4263597"
                     /clone_lib="NCI_CGAP_HN20"
                     /lab_host="DH10B"
                     /note="Organ: normal head/neck tissue; Vector: pAMP1; mRNA
                     made from head/neck tissue, cDNA made by oligo-dT
                     priming. Directionally cloned into UDG sites.
                     Size-selected on  agarose gel, average insert size 300 bp.
                     Primary library.  cDNA Library Preparation: David B.
                     Krizman, Ph.D."
BASE COUNT       11 a      9 c      9 g     20 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 12;  Length 49;
  Best Local Similarity   69.0%;  Pred. No. 3.1e+06;
  Matches   29;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    208 CCATATGAACTTGGCATGGAACATTGTGAGAAATTTGAAATC 249
          || ||||| | ||   | | |||| || ||||  ||||| ||
Db      6 CCTTATGAGCATGTTTTTGTACATCGTTAGAAGCTTGAATTC 47


RESULT 94
BG819402
LOCUS       BG819402                 49 bp    mRNA    linear   EST 22-MAY-2001
DEFINITION  602781632F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4932703
            5', mRNA sequence.
ACCESSION   BG819402
VERSION     BG819402.1  GI:14166989
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10858  row: o  column: 08
            High quality sequence stop: 49.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4932703"
                     /clone_lib="NCI_CGAP_Brn67"
                     /tissue_type="anaplastic oligodendroglioma with 1p/19q
                     loss"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 2.3 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT       29 a      7 c      2 g     11 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 12;  Length 49;
  Best Local Similarity   76.5%;  Pred. No. 3.1e+06;
  Matches   26;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   1762 AAAATCAATCAATGGTGCAAAAAAAAACTTAAAG 1795
           ||||||| |||||    ||||||||||   ||||
Db     16 AAAATCACTCAATTAATCAAAAAAAAAAAAAAG 49


RESULT 95
BJ032647
LOCUS       BJ032647                 51 bp    mRNA    linear   EST 06-DEC-2001
DEFINITION  BJ032647 NIBB Mochii normalized Xenopus neurula library Xenopus
            laevis cDNA clone XL017f01 5', mRNA sequence.
ACCESSION   BJ032647
VERSION     BJ032647.1  GI:17387532
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            ,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XL017f01"
                     /clone_lib="NIBB Mochii normalized Xenopus neurula
                     library"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 15"
                     /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
                     were oligo-dT primed and directionally cloned. Staging
                     according to  Nieuwkoop and Faber. Library is subtracted
                     and was constructed by N. Garrett  and A.M. Zorn,
                     (Wellcome/CRC Institute). "
BASE COUNT       18 a      0 c      2 g     29 t      2 others
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 13;  Length 51;
  Best Local Similarity   72.2%;  Pred. No. 3.1e+06;
  Matches   26;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAATAGA 2143
           | |||||||||| ||||||  |  |||||  || | |
Db     12 TTTTTTTTTTAAATAAATNNAATTTTTTTTAAAAAA 47


RESULT 96
AZ948877/c
LOCUS       AZ948877                 51 bp    DNA     linear   GSS 27-APR-2001
DEFINITION  2M0212P06F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
            clone UUGC2M0212P06 F, DNA sequence.
ACCESSION   AZ948877
VERSION     AZ948877.1  GI:13820104
```

```
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0212  row: P  column: 06
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 51.
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0212P06"
                     /clone_lib="Mouse 10kb plasmid UUGC2M library"
                     /sex="Female"
                     /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (female) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       14 a      4 c      7 g     26 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 17;  Length 51;
  Best Local Similarity   64.0%;  Pred. No. 3.1e+06;
  Matches   32;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy   1273 CCATCTGTACTTGAAAGTGTGAAAGAAAAGTTTTCCTTTGAACCAAAAAT 1322
           ||| |||||||| |||      || || | ||| | | || |  |||| |
Db     50 CCAACTGTACTTAAAAAGAAATAACAATAATTTACTTATGCAAGAAAATT 1


RESULT 97
AL587783/c
LOCUS       AL587783                 52 bp    mRNA    linear   EST 02-MAR-2001
DEFINITION  AL587783 BP Chicken Brain Library Gallus gallus cDNA clone
            ROS063B12, mRNA sequence.
ACCESSION   AL587783
VERSION     AL587783.1  GI:13192817
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Murray,F.
  TITLE     BP Chicken Brain Library
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Frazer Murray
            Dept. Genomics and Bioinformatics
            Roslin Institute
            Roslin, Midlothian, EH25 9PS, UK
            Tel: +44 (0)131 527 4200
            Fax: +44 (0)131 440 0434
            Email: frazer.murray@bbsrc.ac.uk
            GCGGCCGCTTTTTTTTTTTTTTTTT 3' Poly A RNA purchased from Clonetech
            (*6854-
            Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="ROS063B12"
                     /clone_lib="BP Chicken Brain Library"
                     /tissue_type="Brain"
                     /dev_stage="Unknown"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: Not1; Site_2: Sal1; Cloned
                     unidirectionally. Primer: Oligo dT. 5' adaptor sequence:
                     5' TCGACCTCGAG 3' ; 3' adaptor sequence: 5'
                     GCGGCCGCTTTTTTTTTTTTTTTTT 3' Poly A RNA purchased from
                     Clonetech (*6854-1)"
BASE COUNT        9 a      4 c      2 g     37 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 9;  Length 52;
  Best Local Similarity   69.0%;  Pred. No. 3.1e+06;
  Matches   29;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   1761 AAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
           |||||  | | |||||   |||||||||   |||  |||| |
Db     43 AAAAACTATTTAATGGGCAAAAAAAAAAAAAAAAAAAAAAAA 2


RESULT 98
AZ862273
LOCUS       AZ862273                 53 bp    DNA     linear   GSS 21-FEB-2001
DEFINITION  2M0169H19F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0169H19 F, DNA sequence.
ACCESSION   AZ862273
VERSION     AZ862273.1  GI:13059411
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 53)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
```

```
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0169  row: H  column: 19
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 53.
FEATURES              Location/Qualifiers
     source           1. .53
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC2M0169H19"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT       18 a      1 c      31 g      3 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 17;  Length 53;
  Best Local Similarity   64.0%;  Pred. No. 3.1e+06;
  Matches   32;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy     149 AGGATGAGCTGGAGGAGGGGGGTCAGTTAAATGAAAGCATGGACCATGGG 198
           || |||||  |||||||| |||  |   | | |  || ||||    ||||
Db       4 AGCATGAGAGGGAGGAGGAGGGGGAAAGAGAAGGGAGGATGGGGAGTGGG 53


RESULT 99
CNS013L5/c
LOCUS       CNS013L5                  53 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC
            BACN10P12 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL102995
VERSION     AL102995.1  GI:5614606
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 53)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES              Location/Qualifiers
     source           1. .53
                      /organism="Drosophila melanogaster"
                      /db_xref="taxon:7227"
                      /clone="BACN10P12"
                      /clone_lib="DrosBAC"
                      /plasmid="pBeloBAC11"
                      /note="end : T7"
BASE COUNT       46 a      0 c      0 g      4 t      3 others
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 17;  Length 53;
  Best Local Similarity   65.9%;  Pred. No. 3.1e+06;
  Matches   29;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    2111 TTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCA 2154
           ||||||| | |   | | | |||||  :||| ||||||  || |
Db      44 TTTTTTTTTTTTTTTTTTTTTTTTTTTWATATATTTTTTTTTA 1


RESULT 100
BG022522/c
LOCUS       BG022522                  54 bp    mRNA    linear   EST 24-JAN-2001
DEFINITION  daa76h12.x1 Cho Li treated gastrula Xenopus laevis cDNA clone
            IMAGE:4061831 3', mRNA sequence.
ACCESSION   BG022522
VERSION     BG022522.1  GI:12478601
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.
            , Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
            ,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     WashU Xenopus EST project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Sandy Clifton, Ph.D.
            WashU Xenopus EST project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library constructed by C. Hashimoto, Ph.D. in the laboratory of K.
            Cho, Ph.D. DNA Sequencing by: Washington University Genome
            Sequencing Center
             Clone distribution: Xenopus clones from this library are available
            through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
            Seq primer: -40UP from Gibco.
FEATURES              Location/Qualifiers
     source           1. .54
                      /organism="Xenopus laevis"
                      /db_xref="taxon:8355"
                      /clone="IMAGE:4061831"
                      /clone_lib="Cho Li treated gastrula"
                      /tissue_type="gastrula, Li-treated"
                      /lab_host="DH10B"
                      /note="Vector: pBluescript KS+; Site_1: NotI; Site_2:
                      EcoRI; 1st strand was primed with a Not I - oligo(dT)
                      primer, double-stranded cDNA was cloned into the Not I and
                      Eco RI sites of pBluescript KS+. Library was constructed
                      by C. Hashimoto, Ph.D., in the laboratory of K. Cho, Ph.D.
                      (Department of Developmental and Cell Biology, University
                      of California, Irvine)."
BASE COUNT       19 a      2 c      1 g      32 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 12;  Length 54;
```

Best Local Similarity 76.5%; Pred. No. 3.1e+06;
Matches 26; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy   2110 TTTTTTTTATATAAATATATATTTTTCAAATAGA 2143
          ||| ||| | ||||||||| |||||| | | | |
Db     48 TTTGTTTAAAATAAATATAGATTTTTAATAAAAA 15
```

Search completed: April 19, 2003, 10:05:35
Job time : 3729 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: April 19, 2003, 05:16:28 ; Search time 6846 Seconds
(without alignments)
9973.004 Million cell updates/sec

Title: US-09-920-677-3
Perfect score: 2346
Sequence: 1 gcacgaggctgcggcgggtc.....gtaacccagctgcggagcct 2346

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 897812

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	25.8	1.1	59	8	S71517	S71517 (3' flankin
2	25.6	1.1	51	6	AX203959	AX203959 Sequence
3	25	1.1	25	6	AX282921	AX282921 Sequence
c 4	25	1.1	25	6	AX282922	AX282922 Sequence
c 5	24.4	1.0	55	11	AL807848	AL807848 Arabidops
6	24	1.0	30	6	A48425	A48425 Sequence 48
7	23.8	1.0	32	6	A48426	A48426 Sequence 49
8	23.8	1.0	54	6	I08136	I08136 Sequence 1
9	23.6	1.0	50	6	AX164846	AX164846 Sequence
10	23.4	1.0	43	6	AX484399	AX484399 Sequence
11	23.4	1.0	54	6	AR002541	AR002541 Sequence
12	23.4	1.0	58	6	A51709	A51709 Sequence 15
13	23.4	1.0	58	6	AR167588	AR167588 Sequence
14	23.4	1.0	58	6	AR178298	AR178298 Sequence
15	23.4	1.0	58	6	AX323380	AX323380 Sequence
c 16	23	1.0	58	6	A08890	A08890 H.sapiens (
17	22.8	1.0	48	6	AX003947	AX003947 Sequence
18	22.8	1.0	48	6	AX021566	AX021566 Sequence
19	22.8	1.0	48	6	AX021576	AX021576 Sequence
20	22.8	1.0	48	6	AX021632	AX021632 Sequence
c 21	22.6	1.0	51	6	AX115973	AX115973 Sequence
22	22.6	1.0	54	6	AR089685	AR089685 Sequence
23	22.6	1.0	54	6	AX317217	AX317217 Sequence
24	22.4	1.0	44	6	AR032451	AR032451 Sequence
25	22.4	1.0	44	6	AR209115	AR209115 Sequence
26	22.4	1.0	44	6	I29191	I29191 Sequence 63
27	22.4	1.0	44	6	I90865	I90865 Sequence 63
28	22.2	0.9	45	6	AR032600	AR032600 Sequence
29	22.2	0.9	45	6	AR209264	AR209264 Sequence
30	22.2	0.9	45	6	I29340	I29340 Sequence 21
31	22.2	0.9	45	6	I91014	I91014 Sequence 21
c 32	22.2	0.9	47	6	AX114334	AX114334 Sequence
c 33	22	0.9	33	6	AX172247	AX172247 Sequence
34	22	0.9	42	6	AR152463	AR152463 Sequence
c 35	22	0.9	50	6	AX164855	AX164855 Sequence
c 36	22	0.9	58	6	A51710	A51710 Sequence 16
c 37	22	0.9	58	6	AR167589	AR167589 Sequence
c 38	22	0.9	58	6	AR178299	AR178299 Sequence
c 39	22	0.9	58	6	AX323381	AX323381 Sequence
40	22	0.9	60	6	BD008815	BD008815 Adhesin b
c 41	21.8	0.9	50	6	AR117934	AR117934 Sequence
c 42	21.8	0.9	50	6	AR117952	AR117952 Sequence
43	21.8	0.9	51	6	AX160479	AX160479 Sequence
c 44	21.6	0.9	51	6	AX160117	AX160117 Sequence
c 45	21.6	0.9	51	6	AX160118	AX160118 Sequence
c 46	21.6	0.9	54	9	AF079025	AF079025 Homo sapi
c 47	21.6	0.9	55	6	AX484751	AX484751 Sequence
48	21.4	0.9	31	6	AX248729	AX248729 Sequence
49	21.4	0.9	48	6	AR064025	AR064025 Sequence
50	21.4	0.9	48	6	AR178317	AR178317 Sequence
51	21.4	0.9	48	6	AX323399	AX323399 Sequence
c 52	21.4	0.9	51	6	AX162478	AX162478 Sequence
53	21.4	0.9	55	6	I29928	I29928 Sequence 41
54	21.4	0.9	60	6	I66359	I66359 Sequence 18
c 55	21.2	0.9	50	6	AX160974	AX160974 Sequence
c 56	21.2	0.9	51	6	AX157225	AX157225 Sequence
c 57	21.2	0.9	51	6	AX160455	AX160455 Sequence
58	21.2	0.9	54	6	AR192747	AR192747 Sequence
c 59	21.2	0.9	54	9	AF254565	AF254565 Homo sapi
c 60	21.2	0.9	55	6	A43625	A43625 Sequence 15
61	21.2	0.9	56	6	A80902	A80902 Sequence 13
62	21.2	0.9	56	6	A97341	A97341 Sequence 13
63	21.2	0.9	56	6	A97358	A97358 Sequence 13
64	21.2	0.9	56	6	AR183071	AR183071 Sequence
65	21	0.9	38	6	E07494	E07494 Synthetic D

c	66	21	0.9	40	6	AX299737	AX299737 Sequence
	67	21	0.9	51	6	AX160105	AX160105 Sequence
	68	21	0.9	56	10	MUSBMP24H	D45009 Mouse gene
	69	21	0.9	59	6	AR209950	AR209950 Sequence
	70	21	0.9	59	6	AX253387	AX253387 Sequence
	71	21	0.9	59	6	AX254767	AX254767 Sequence
	72	21	0.9	60	6	AR073297	AR073297 Sequence
c	73	21	0.9	60	8	YSCMTP191	J01506 Yeast (S.ce
	74	20.8	0.9	57	3	AF177252	AF177252 Bodo salt
c	75	20.6	0.9	28	6	A48423	A48423 Sequence 46
c	76	20.6	0.9	28	6	A48424	A48424 Sequence 47
	77	20.6	0.9	50	6	A51711	A51711 Sequence 17
	78	20.6	0.9	50	6	AR167590	AR167590 Sequence
	79	20.6	0.9	50	6	AR178300	AR178300 Sequence
	80	20.6	0.9	50	6	AX165034	AX165034 Sequence
	81	20.6	0.9	50	6	AX323382	AX323382 Sequence
	82	20.6	0.9	51	6	AR068824	AR068824 Sequence
	83	20.6	0.9	51	6	AR122557	AR122557 Sequence
c	84	20.6	0.9	51	6	AX157089	AX157089 Sequence
c	85	20.6	0.9	51	6	AX157090	AX157090 Sequence
	86	20.6	0.9	51	6	AX158178	AX158178 Sequence
c	87	20.6	0.9	51	6	AX161887	AX161887 Sequence
c	88	20.6	0.9	51	6	AX161888	AX161888 Sequence
c	89	20.6	0.9	51	6	AX162091	AX162091 Sequence
c	90	20.6	0.9	51	6	AX162092	AX162092 Sequence
c	91	20.6	0.9	53	6	AX404671	AX404671 Sequence
c	92	20.6	0.9	54	10	AF328254	AF328254 Mus muscu
	93	20.6	0.9	60	6	E10890	E10890 PCR primer
	94	20.4	0.9	51	6	AX117117	AX117117 Sequence
	95	20.4	0.9	57	8	YSCMTP182	J01505 Yeast (S.ce
	96	20.2	0.9	39	6	A08918	A08918 H.sapiens (
c	97	20.2	0.9	45	9	HSU26975	U26975 Human isola
	98	20.2	0.9	48	6	AR032407	AR032407 Sequence
	99	20.2	0.9	48	6	AR209071	AR209071 Sequence
	100	20.2	0.9	48	6	I29147	I29147 Sequence 19

ALIGNMENTS

```
RESULT 1
S71517
LOCUS       S71517                    59 bp    DNA     linear   PLN 28-OCT-1994
DEFINITION  {3' flanking DNA of rpl2 exon 2 homolog} [Pisum sativum, Genomic,
            59 nt].
ACCESSION   S71517
VERSION     S71517.1  GI:560660
KEYWORDS    .
SOURCE      Pisum sativum.
  ORGANISM  Pisum sativum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae;
            Pisum.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Kotliarova,E.G., Denisenko,Iu.V. and Tarasov,V.A.
  TITLE     Analysis of nucleotide sequences of 0.7 kb fragments from mixed DNA
            from Arabidopsis thaliana and Pisum sativum
  JOURNAL   Genetika 30 (4), 565-569 (1994)
  MEDLINE   94320766
   PUBMED   8045406
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 150347] from the original journal article.
            This sequence comes from Fig. 2.
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Pisum sativum"
                     /db_xref="taxon:3888"
BASE COUNT       16 a     14 c      8 g     21 t
ORIGIN

  Query Match             1.1%;  Score 25.8;  DB 8;  Length 59;
  Best Local Similarity   67.9%;  Pred. No. 2.3e+05;
  Matches   36;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   2134 TTCAAATAGATTTTTGATTCAGCTCATTATGAAAAACATCCCAAACTTTAAAA 2186
           || || |||||| ||  ||  | |||||   ||   | |||||||||| || ||
Db      3 TTGAAGTAGATTCTTCTTTTTGATCATTCAAAACCCCTTCCCAAACTGTACAA 55


RESULT 2
AX203959
LOCUS       AX203959                  51 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 65 from Patent WO0148245.
ACCESSION   AX203959
VERSION     AX203959.1  GI:15393422
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0148245-A 65 05-JUL-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     variation       26
                     /note="single nucleotide polymorphism
                     Accession number cg43969763"
BASE COUNT       15 a     11 c      6 g     19 t
ORIGIN

  Query Match             1.1%;  Score 25.6;  DB 6;  Length 51;
  Best Local Similarity   70.8%;  Pred. No. 2.5e+05;
  Matches   34;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    665 GGATCATCTACAGAGACCTGAAGCCGGAGAATATCATGCTTAATCACC 712
          | |||||  || | || || ||||| || |||||| | |||  | |||
Db      4 GTATCATTCACTGTGATCTAAAGCCTGAAAATATCCTTCTTTGTAACC 51


RESULT 3
AX282921
LOCUS       AX282921                  25 bp    DNA     linear   PAT 02-NOV-2001
DEFINITION  Sequence 11 from Patent WO0177338.
ACCESSION   AX282921
VERSION     AX282921.1  GI:16609890
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Plowman,G., Whyte,D., Manning,G., Sudarsanam,S., Martinez,R. and
            Caenepeel,S.
  TITLE     Human protein kinases and protein kinase-like enzymes
  JOURNAL   Patent: WO 0177338-A 11 18-OCT-2001;
            Sugen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Primer"
BASE COUNT       10 a      2 c      6 g      7 t
ORIGIN

  Query Match             1.1%;  Score 25;  DB 6;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+05;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    560 GAGAACTATTTATGCAGTTAGAAAG 584
          |||||||||||||||||||||||||
```

```
Db      1 GAGAACTATTTATGCAGTTAGAAAG 25


RESULT 4
AX282922/c
LOCUS       AX282922                  25 bp    DNA     linear   PAT 02-NOV-2001
DEFINITION  Sequence 12 from Patent WO0177338.
ACCESSION   AX282922
VERSION     AX282922.1  GI:16609891
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Plowman,G., Whyte,D., Manning,G., Sudarsanam,S., Martinez,R. and
            Caenepeel,S.
  TITLE     Human protein kinases and protein kinase-like enzymes
  JOURNAL   Patent: WO 0177338-A 12 18-OCT-2001;
            Sugen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Primer"
BASE COUNT        6 a      6 c      5 g      8 t
ORIGIN

  Query Match             1.1%;  Score 25;  DB 6;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+05;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1087 CACATTAACTGGGAAGAACTTCTGG 1111
           |||||||||||||||||||||||||
Db      25 CACATTAACTGGGAAGAACTTCTGG 1


RESULT 5
AL807848/c
LOCUS       AL807848                  55 bp    DNA     linear   STS 03-JUL-2002
DEFINITION  Arabidopsis thaliana transposon insertion STS SM_3.2611, sequence
            tagged site.
ACCESSION   AL807848
VERSION     AL807848.1  GI:21690148
KEYWORDS    STS; STS, sequence tagged site.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B.,
            Murphy,G., Langham,S., LeGrys,C., Jones,J.D.G. and Bevan,M.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 55)
  AUTHORS   Clarke,J.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUL-2002) Clarke J.H., John Innes Centre, Colney
            Lane, Norwich, NR4 7UJ, UK
COMMENT     AT denotes an activation tag dissociation transposon within a
            single line, ET an enhancer trap dissociation transposon, GT a gene
            trap dissociation transposon, MT a mis-expression enhancer trap
            dissociation transposon, SM a defective suppressor mutator
            transposon. _3 denotes a sequence derived from the 3'end of the
            transposon, _5 denotes a sequence derived from the 5'end of the
            transposon BBSRC GARNet, ATIS project
            On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock
            code: N109576.
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Arabidopsis thaliana"
                     /variety="Landsberg erecta"
                     /db_xref="taxon:3702"
                     /clone="AL035605"
     STS             1. .55
                     /standard_name="SM_3.2611"
BASE COUNT       27 a      1 c      2 g     25 t
ORIGIN

  Query Match             1.0%;  Score 24.4;  DB 11;  Length 55;
  Best Local Similarity   73.8%;  Pred. No. 4.8e+05;
  Matches   31;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy    2111 TTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATT 2152
           | | | |||||| |||||||| | | |||   |||||| |||
Db      55 TCTATATATATATATATATATATATAAAACTAATTTTTTATT 14


RESULT 6
A48425
LOCUS       A48425                    30 bp    DNA     linear   PAT 07-MAR-1997
DEFINITION  Sequence 48 from Patent WO9603501.
ACCESSION   A48425
VERSION     A48425.1  GI:2302216
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Chaudhuri,B., Stephan,C. and Fuerst,P.
  TITLE     DUAL HYBRID SYSTEM
  JOURNAL   Patent: WO 9603501-A 48 08-FEB-1996;
            CIBA GEIGY AG (CH)
COMMENT     Other publication AU 2983295 960222.
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     mat_peptide     10. .30
                     /note="TRUNCATED P70S6K DELTA N GENE"
                     /function="PCR FRAGMENT"
BASE COUNT        7 a      5 c      8 g     10 t
ORIGIN

  Query Match             1.0%;  Score 24;  DB 6;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 5.6e+05;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     289 CCAGAATGTTTTGAGCTACTTCGG 312
           ||||||||||||||||||||||||
Db       7 CCAGAATGTTTTGAGCTACTTCGG 30


RESULT 7
A48426
LOCUS       A48426                    32 bp    DNA     linear   PAT 07-MAR-1997
DEFINITION  Sequence 49 from Patent WO9603501.
ACCESSION   A48426
VERSION     A48426.1  GI:2302217
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 32)
  AUTHORS   Chaudhuri,B., Stephan,C. and Fuerst,P.
  TITLE     DUAL HYBRID SYSTEM
  JOURNAL   Patent: WO 9603501-A 49 08-FEB-1996;
            CIBA GEIGY AG (CH)
COMMENT     Other publication AU 2983295 960222.
FEATURES             Location/Qualifiers
     source          1. .32
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     mat_peptide     10. .32
                     /note="SENSE STRAND OF THE
                     312 BP AUTOINHIBITORY SEQUENCE FROM P70S6K GENE"
                     /function="PCR FRAGMENT"
```

```
BASE COUNT       12 a      4 c      5 g     11 t
ORIGIN

  Query Match             1.0%;  Score 23.8;  DB 6;  Length 32;
  Best Local Similarity   92.6%;  Pred. No. 6.3e+05;
  Matches   25;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy   1290 TGTGAAAGAAAAGTTTTCCTTTGAACC 1316
          | ||||||||||||||||| ||||||||
Db      6 TCTGAAAGAAAAGTTTTCTTTTGAACC 32


RESULT 8
I08136
LOCUS       I08136                   54 bp    DNA     linear   PAT 02-DEC-1994
DEFINITION  Sequence 1 from Patent EP 0343783.
ACCESSION   I08136
VERSION     I08136.1  GI:589151
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Barsoum,J.G.
  TITLE     Method for producing cells containing stably integrated foreign DNA
            at a high copy number, the cells produced by this method, and the
            use of these cells to produce the polypeptides coded for by the
            foreign DNA
  JOURNAL   Patent: EP 0343783-A2 1 29-NOV-1989;
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="unknown"
BASE COUNT       17 a      0 c      0 g     37 t
ORIGIN

  Query Match             1.0%;  Score 23.8;  DB 6;  Length 54;
  Best Local Similarity   66.7%;  Pred. No. 6.6e+05;
  Matches   34;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   2113 TTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCAGCTCATTAT 2163
          | ||||| || |||| ||| |||   | | ||| || ||| |  |  ||||
Db      2 TATTTATTTATATATTTATATTTTTTAAATATTATTTATTTATTTATTTAT 52


RESULT 9
AX164846
LOCUS       AX164846                 50 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 41 from Patent WO0138586.
ACCESSION   AX164846
VERSION     AX164846.1  GI:14545675
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0138586-A 41 31-MAY-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    25. .26
                     /note="Nucleotide deleted between bases 25 and 26
                     Accession number cg43988092"
     variation       26
                     /note="single nucleotide polymorphism"
BASE COUNT       22 a      5 c      3 g     20 t
ORIGIN

  Query Match             1.0%;  Score 23.6;  DB 6;  Length 50;
  Best Local Similarity   76.3%;  Pred. No. 7.3e+05;
  Matches   29;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy   2117 TATATAAATATATATTTTTCAAATAGATTTTTGATTCA 2154
          |||||| |||||| || | |||| | || ||| || ||
Db     13 TATATATATATATTTTATCCAAAAATATGTTTTATACA 50


RESULT 10
AX484399
LOCUS       AX484399                 43 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 1699 from Patent WO02053728.
ACCESSION   AX484399
VERSION     AX484399.1  GI:22318751
KEYWORDS    .
SOURCE      Candida albicans.
  ORGANISM  Candida albicans
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 02053728-A 1699 11-JUL-2002;
            Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="Candida albicans"
                     /db_xref="taxon:5476"
BASE COUNT       14 a      4 c      3 g     22 t
ORIGIN

  Query Match             1.0%;  Score 23.4;  DB 6;  Length 43;
  Best Local Similarity   81.8%;  Pred. No. 8e+05;
  Matches   27;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy   2116 TTATATAAATATATATTTTTCAAATAGATTTTT 2148
          ||||| | |||||||||| || |||| | |||||
Db      8 TTATACATATATATATTGTTTAAATTGTTTTTT 40


RESULT 11
AR002541
LOCUS       AR002541                 54 bp    DNA     linear   PAT 04-DEC-1998
DEFINITION  Sequence 72 from patent US 5741696.
ACCESSION   AR002541
VERSION     AR002541.1  GI:3964095
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Cochran,M.D. and Chiang,C.H.
  TITLE     Recombinant equine herpesviruses
  JOURNAL   Patent: US 5741696-A 72 21-APR-1998;
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="unknown"
BASE COUNT       19 a      11 c      8 g     16 t
ORIGIN

  Query Match             1.0%;  Score 23.4;  DB 6;  Length 54;
  Best Local Similarity   67.3%;  Pred. No. 8.2e+05;
  Matches   33;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    760 CATGATGGAACAGTCACACACACATTTTGTGGAACAATAGAATACATGG 808
          | | ||| | ||  |||| ||   ||  |||  || ||||||||||| ||
Db      1 CCTTATGTATCATACACATACGATTTAGGTGACACTATAGAATACACGG 49


RESULT 12
A51709
LOCUS       A51709                   58 bp    DNA     linear   PAT 10-MAR-1997
```

```
DEFINITION  Sequence 15 from Patent WO9618744.
ACCESSION   A51709
VERSION     A51709.1  GI:2304513
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Crouzet,J., Scherman,D. and Wils,P.
  TITLE     PURIFICATION OF A TRIPLE HELIX FORMATION WITH AN IMMOBILIZED
            OLIGONUCLEOTIDE
  JOURNAL   Patent: WO 9618744-A 15 20-JUN-1996;
            RHONE POULENC RORER SA (FR)
COMMENT     Other publication AU 4178996 960703
            Other publication FR 2728264 960621.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       18 a      2 c     37 g      1 t
ORIGIN

  Query Match             1.0%;  Score 23.4;  DB 6;  Length 58;
  Best Local Similarity   63.2%;  Pred. No. 8.2e+05;
  Matches   36;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy    118 GACCTGGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAG 174
          || | |||  ||   ||||| | ||     ||||| |||  |||||||| ||   ||
Db      1 GATCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 57


RESULT 13
AR167588
LOCUS       AR167588                 58 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 15 from patent US 6287762.
ACCESSION   AR167588
VERSION     AR167588.1  GI:17903377
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Crouzet,J., Scherman,D. and Wils,P.
  TITLE     Purification of a triple helix formation with an immobilized
            oligonucleotide
  JOURNAL   Patent: US 6287762-A 15 11-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="unknown"
BASE COUNT       18 a      2 c     37 g      1 t
ORIGIN

  Query Match             1.0%;  Score 23.4;  DB 6;  Length 58;
  Best Local Similarity   63.2%;  Pred. No. 8.2e+05;
  Matches   36;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy    118 GACCTGGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAG 174
          || | |||  ||   ||||| | ||     ||||| |||  |||||||| ||   ||
Db      1 GATCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 57


RESULT 14
AR178298
LOCUS       AR178298                 58 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 15 from patent US 6319672.
ACCESSION   AR178298
VERSION     AR178298.1  GI:20219436
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Crouzet,J., Scherman,D., Wils,P., Blanche,F. and Cameron,B.
  TITLE     Purification of a triple helix formation with an immobilized
            oligonucleotide
  JOURNAL   Patent: US 6319672-A 15 20-NOV-2001;
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="unknown"
BASE COUNT       18 a      2 c     37 g      1 t
ORIGIN

  Query Match             1.0%;  Score 23.4;  DB 6;  Length 58;
  Best Local Similarity   63.2%;  Pred. No. 8.2e+05;
  Matches   36;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy    118 GACCTGGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAG 174
          || | |||  ||   ||||| | ||     ||||| |||  |||||||| ||   ||
Db      1 GATCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 57


RESULT 15
AX323380
LOCUS       AX323380                 58 bp    DNA     linear   PAT 07-JAN-2002
DEFINITION  Sequence 15 from Patent WO0192511.
ACCESSION   AX323380
VERSION     AX323380.1  GI:18094142
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Crouzet,J., Scherman,D., Wils,P., Blanche,F. and Cameron,B.
  TITLE     Purification of a triple helix formation with an immobilized
            oligonucleotide
  JOURNAL   Patent: WO 0192511-A 15 06-DEC-2001;
            Aventis Pharma (FR)
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic oligonucleotide"
BASE COUNT       18 a      2 c     37 g      1 t
ORIGIN

  Query Match             1.0%;  Score 23.4;  DB 6;  Length 58;
  Best Local Similarity   63.2%;  Pred. No. 8.2e+05;
  Matches   36;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy    118 GACCTGGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAG 174
          || | |||  ||   ||||| | ||     ||||| |||  |||||||| ||   ||
Db      1 GATCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 57


RESULT 16
A08890/c
LOCUS       A08890                   58 bp    DNA     linear   PAT 02-SEP-1993
DEFINITION  H.sapiens (haplotype 1, allele MS32, isolate French, serial number
            6) minisatellite sequence.
ACCESSION   A08890
VERSION     A08890.1  GI:411812
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Jeffreys,A.J.
  TITLE     Extended nucleotide sequences
  JOURNAL   Patent: EP 0370719-A 70 30-MAY-1990;
            IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       34 a      0 c      0 g     24 t
```

```
ORIGIN

  Query Match           1.0%;  Score 23;  DB 6;  Length 58;
  Best Local Similarity   68.1%;  Pred. No. 1e+06;
  Matches   32;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2104 TGAATCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGA 2150
          | ||| |||||||| | |  |||||| | |   ||||| | | || | |
Db     52 TTAATTTTTTTTTTTTTTTTATATATTTATAATAAATAAAATATTTA 6


RESULT 17
AX003947
LOCUS       AX003947                 48 bp    DNA     linear   PAT 24-AUG-2000
DEFINITION  Sequence 7 from Patent WO9923249.
ACCESSION   AX003947
VERSION     AX003947.1  GI:9927607
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Kessler,C. and Bartl,K.
  TITLE     Specific and sensitive method for detecting nucleic acids
  JOURNAL   Patent: WO 9923249-A 7 14-MAY-1999;
            KESSLER CHRISTOPH (DE); BARTL KNUT (DE)
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        9 a     17 c     14 g      8 t
ORIGIN

  Query Match           1.0%;  Score 22.8;  DB 6;  Length 48;
  Best Local Similarity   71.4%;  Pred. No. 1.1e+06;
  Matches   30;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    874 ATGTATGACATGCTGACTGGAGCACCCCCATTCACTGGGGAG 915
          |||| || | ||| | ||  || |||||| || | || |||
Db      4 ATGTGTGTCGTGCAGCCTCCAGGACCCCCACTCCCGGGAGAG 45


RESULT 18
AX021566
LOCUS       AX021566                 48 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 4 from Patent WO9924606.
ACCESSION   AX021566
VERSION     AX021566.1  GI:10044850
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Kessler,C., Bartl,K., Haberhausen,G. and Orum,H.
  TITLE     Specific and sensitive nucleic acid detection method
  JOURNAL   Patent: WO 9924606-A 4 20-MAY-1999;
            KESSLER CHRISTOPH (DE); BARTL KNUT (DE); HABERHAUSEN GERD (DE);
            ROCHE DIAGNOSTICS GMBH (DE); ORUM HENRIK (DK)
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        9 a     17 c     14 g      8 t
ORIGIN

  Query Match           1.0%;  Score 22.8;  DB 6;  Length 48;
  Best Local Similarity   71.4%;  Pred. No. 1.1e+06;
  Matches   30;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    874 ATGTATGACATGCTGACTGGAGCACCCCCATTCACTGGGGAG 915
          |||| || | ||| | ||  || |||||| || | || |||
Db      4 ATGTGTGTCGTGCAGCCTCCAGGACCCCCACTCCCGGGAGAG 45


RESULT 19
AX021576
LOCUS       AX021576                 48 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 14 from Patent WO9924606.
ACCESSION   AX021576
VERSION     AX021576.1  GI:10044860
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Kessler,C., Bartl,K., Haberhausen,G. and Orum,H.
  TITLE     Specific and sensitive nucleic acid detection method
  JOURNAL   Patent: WO 9924606-A 14 20-MAY-1999;
            KESSLER CHRISTOPH (DE); BARTL KNUT (DE); HABERHAUSEN GERD (DE);
            ROCHE DIAGNOSTICS GMBH (DE); ORUM HENRIK (DK)
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="HCV_MCR"
BASE COUNT        9 a     17 c     14 g      8 t
ORIGIN

  Query Match           1.0%;  Score 22.8;  DB 6;  Length 48;
  Best Local Similarity   71.4%;  Pred. No. 1.1e+06;
  Matches   30;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    874 ATGTATGACATGCTGACTGGAGCACCCCCATTCACTGGGGAG 915
          |||| || | ||| | ||  || |||||| || | || |||
Db      4 ATGTGTGTCGTGCAGCCTCCAGGACCCCCACTCCCGGGAGAG 45


RESULT 20
AX021632
LOCUS       AX021632                 48 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 11 from Patent WO9923250.
ACCESSION   AX021632
VERSION     AX021632.1  GI:10044915
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Kessler,C., Bartl,K., Haberhausen,G. and Orum,H.
  TITLE     Specific and sensitive method for detecting nucleic acids
  JOURNAL   Patent: WO 9923250-A 11 14-MAY-1999;
            KESSLER CHRISTOPH (DE); BARTL KNUT (DE); HABERHAUSEN GERD (DE);
            ROCHE DIAGNOSTICS GMBH (DE); ORUM HENRIK (DK)
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="HCV_MCR"
BASE COUNT        9 a     17 c     14 g      8 t
ORIGIN

  Query Match           1.0%;  Score 22.8;  DB 6;  Length 48;
  Best Local Similarity   71.4%;  Pred. No. 1.1e+06;
  Matches   30;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    874 ATGTATGACATGCTGACTGGAGCACCCCCATTCACTGGGGAG 915
          |||| || | ||| | ||  || |||||| || | || |||
Db      4 ATGTGTGTCGTGCAGCCTCCAGGACCCCCACTCCCGGGAGAG 45


RESULT 21
AX115973/c
LOCUS       AX115973                 51 bp    DNA     linear   PAT 11-MAY-2001
DEFINITION  Sequence 1096 from Patent WO0129262.
```

```
ACCESSION   AX115973
VERSION     AX115973.1  GI:14032915
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Picoult-Newburg,L. and Pohl,M.
  TITLE     Genotyping reagents, kits and methods of use thereof
  JOURNAL   Patent: WO 0129262-A 1096 26-APR-2001;
            Orchid BioSciences, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       13 a      9 c      8 g     20 t      1 others
ORIGIN

  Query Match             1.0%;  Score 22.6;  DB 6;  Length 51;
  Best Local Similarity   68.9%;  Pred. No. 1.2e+06;
  Matches   31;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    406 AAGGCAATGATAGTAAGAAATGCTAAAGATACAGCTCATACAAAA 450
          | |||||||||| | ||| |   || ||||| | ||  |  ||||
Db     51 ATGGCAATGATATTCAGACACATTACAGATATATCTGCTGGAAAA 7


RESULT 22
AR089685
LOCUS       AR089685                 54 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 162 from patent US 5994069.
ACCESSION   AR089685
VERSION     AR089685.1  GI:10016440
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Hall,J.G., Lyamichev,V.I., Mast,A.L. and Brow,M.Ann.D.
  TITLE     Detection of nucleic acids by multiple sequential invasive
            cleavages
  JOURNAL   Patent: US 5994069-A 162 30-NOV-1999;
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="unknown"
BASE COUNT       14 a      7 c     25 g      8 t
ORIGIN

  Query Match             1.0%;  Score 22.6;  DB 6;  Length 54;
  Best Local Similarity   75.7%;  Pred. No. 1.3e+06;
  Matches   28;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy    128 AGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGA 164
          ||  |||||| | |||  || |||| |||| ||||||
Db      4 AGGAAGAGGAGGAGGGTGCTCAGGAGGAGCGGGAGGA 40


RESULT 23
AX317217
LOCUS       AX317217                 54 bp    DNA     linear   PAT 14-DEC-2001
DEFINITION  Sequence 220 from Patent WO0190337.
ACCESSION   AX317217
VERSION     AX317217.1  GI:17900206
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Allawi,H., Bartholomay,C.T., Chehak,L., Curtis,M.L., Eis,P.S.,
            Hall,J.G., Ip,H.S., Kaiser,M., Kwiatkowski,R.W., Lukowiak,A.A.,
            Lyamichev,V., Ma,W., Olson-Munoz,M.C., Olson,S.M., Schaefer,J.J.,
            Skrzypczynski,Z., Takova,T.Y., Vedvik,K.L. and Lyamichev,N.E.
  TITLE     Detection of rna
  JOURNAL   Patent: WO 0190337-A 220 29-NOV-2001;
            THIRD WAVE TECHNOLOGIES, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT       14 a      7 c     25 g      8 t
ORIGIN

  Query Match             1.0%;  Score 22.6;  DB 6;  Length 54;
  Best Local Similarity   75.7%;  Pred. No. 1.3e+06;
  Matches   28;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy    128 AGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGA 164
          ||  |||||| | |||  || |||| |||| ||||||
Db      4 AGGAAGAGGAGGAGGGTGCTCAGGAGGAGCGGGAGGA 40


RESULT 24
AR032451
LOCUS       AR032451                 44 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 63 from patent US 5869241.
ACCESSION   AR032451
VERSION     AR032451.1  GI:5948056
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Method of determining DNA sequence preference of a DNA-binding
            molecule
  JOURNAL   Patent: US 5869241-A 63 09-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="unknown"
BASE COUNT        6 a     15 c     20 g      3 t
ORIGIN

  Query Match             1.0%;  Score 22.4;  DB 6;  Length 44;
  Best Local Similarity   72.5%;  Pred. No. 1.4e+06;
  Matches   29;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      9 CTGCGGCGGGTCCGGGCCCATGAGGCGACGAAGGAGGCGG 48
          || | ||||   |||||  |||||| | || ||| |||||
Db      4 CTCCCGCGGACGCGGGCATATGAGGAGGCGGAGGCGGCGG 43


RESULT 25
AR209115
LOCUS       AR209115                 44 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 63 from patent US 6384208.
ACCESSION   AR209115
VERSION     AR209115.1  GI:21510447
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Sequence directed DNA binding molecules compositions and methods
  JOURNAL   Patent: US 6384208-A 63 07-MAY-2002;
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="unknown"
BASE COUNT        6 a     15 c     20 g      3 t
ORIGIN

  Query Match             1.0%;  Score 22.4;  DB 6;  Length 44;
  Best Local Similarity   72.5%;  Pred. No. 1.4e+06;
  Matches   29;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;
```

```
Qy      9 CTGCGGCGGGTCCGGGCCCATGAGGCGACGAAGGAGGCGG 48
          || | ||||   |||||  |||||| | || ||| |||||
Db      4 CTCCCGCGGACGCGGGCATATGAGGAGGCGGAGGCGGCGG 43


RESULT 26
I29191
LOCUS       I29191                  44 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 63 from patent US 5578444.
ACCESSION   I29191
VERSION     I29191.1  GI:1819982
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Sequence-directed DNA-binding molecules compositions and methods
  JOURNAL   Patent: US 5578444-A 63 26-NOV-1996;
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="unknown"
BASE COUNT        6 a     15 c     20 g      3 t
ORIGIN

  Query Match             1.0%;  Score 22.4;  DB 6;  Length 44;
  Best Local Similarity   72.5%;  Pred. No. 1.4e+06;
  Matches   29;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      9 CTGCGGCGGGTCCGGGCCCATGAGGCGACGAAGGAGGCGG 48
          || | ||||   |||||  |||||| | || ||| |||||
Db      4 CTCCCGCGGACGCGGGCATATGAGGAGGCGGAGGCGGCGG 43


RESULT 27
I90865
LOCUS       I90865                  44 bp    DNA     linear   PAT 01-DEC-1998
DEFINITION  Sequence 63 from patent US 5726014.
ACCESSION   I90865
VERSION     I90865.1  GI:3935335
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turin,L.M.
  TITLE     Screening assay for the detection of DNA-binding molecules
  JOURNAL   Patent: US 5726014-A 63 10-MAR-1998;
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="unknown"
BASE COUNT        6 a     15 c     20 g      3 t
ORIGIN

  Query Match             1.0%;  Score 22.4;  DB 6;  Length 44;
  Best Local Similarity   72.5%;  Pred. No. 1.4e+06;
  Matches   29;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      9 CTGCGGCGGGTCCGGGCCCATGAGGCGACGAAGGAGGCGG 48
          || | ||||   |||||  |||||| | || ||| |||||
Db      4 CTCCCGCGGACGCGGGCATATGAGGAGGCGGAGGCGGCGG 43


RESULT 28
AR032600
LOCUS       AR032600                45 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 212 from patent US 5869241.
ACCESSION   AR032600
VERSION     AR032600.1  GI:5948205
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Method of determining DNA sequence preference of a DNA-binding
            molecule
  JOURNAL   Patent: US 5869241-A 212 09-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="unknown"
BASE COUNT       13 a     10 c      7 g     15 t
ORIGIN

  Query Match             0.9%;  Score 22.2;  DB 6;  Length 45;
  Best Local Similarity   77.1%;  Pred. No. 1.5e+06;
  Matches   27;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    933 TGACAAAATCCTCAAATGTAAACTCAATTTGCCTC 967
          | || |||   || ||||||||||||| ||| ||||
Db      9 TAACGAAAATTTCCAATGTAAACTCATTTTCCCTC 43


RESULT 29
AR209264
LOCUS       AR209264                45 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 212 from patent US 6384208.
ACCESSION   AR209264
VERSION     AR209264.1  GI:21510637
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Sequence directed DNA binding molecules compositions and methods
  JOURNAL   Patent: US 6384208-A 212 07-MAY-2002;
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="unknown"
BASE COUNT       13 a     10 c      7 g     15 t
ORIGIN

  Query Match             0.9%;  Score 22.2;  DB 6;  Length 45;
  Best Local Similarity   77.1%;  Pred. No. 1.5e+06;
  Matches   27;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    933 TGACAAAATCCTCAAATGTAAACTCAATTTGCCTC 967
          | || |||   || ||||||||||||| ||| ||||
Db      9 TAACGAAAATTTCCAATGTAAACTCATTTTCCCTC 43


RESULT 30
I29340
LOCUS       I29340                  45 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 212 from patent US 5578444.
ACCESSION   I29340
VERSION     I29340.1  GI:1820131
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Sequence-directed DNA-binding molecules compositions and methods
  JOURNAL   Patent: US 5578444-A 212 26-NOV-1996;
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="unknown"
BASE COUNT       13 a     10 c      7 g     15 t
ORIGIN

  Query Match             0.9%;  Score 22.2;  DB 6;  Length 45;
  Best Local Similarity   77.1%;  Pred. No. 1.5e+06;
  Matches   27;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;
```

```
Qy    933 TGACAAAATCCTCAAATGTAAACTCAATTTGCCTC 967
          | || |||   || ||||||||||| ||| ||||
Db      9 TAACGAAAATTTCCAATGTAAACTCATTTTCCCTC 43


RESULT 31
I91014
LOCUS       I91014                    45 bp    DNA     linear   PAT 01-DEC-1998
DEFINITION  Sequence 212 from patent US 5726014.
ACCESSION   I91014
VERSION     I91014.1  GI:3935484
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turin,L.M.
  TITLE     Screening assay for the detection of DNA-binding molecules
  JOURNAL   Patent: US 5726014-A 212 10-MAR-1998;
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="unknown"
BASE COUNT       13 a     10 c      7 g     15 t
ORIGIN

  Query Match             0.9%;  Score 22.2;  DB 6;  Length 45;
  Best Local Similarity   77.1%;  Pred. No. 1.5e+06;
  Matches   27;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    933 TGACAAAATCCTCAAATGTAAACTCAATTTGCCTC 967
          | || |||   || ||||||||||| ||| ||||
Db      9 TAACGAAAATTTCCAATGTAAACTCATTTTCCCTC 43


RESULT 32
AX114334/c
LOCUS       AX114334                  47 bp    DNA     linear   PAT 11-MAY-2001
DEFINITION  Sequence 3 from Patent WO0129257.
ACCESSION   AX114334
VERSION     AX114334.1  GI:14031298
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Schork,N. and Skierczynski,B.
  TITLE     Methods of genetic cluster analysis and use thereof
  JOURNAL   Patent: WO 0129257-A 3 26-APR-2001;
            GENSET (FR)
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     variation       24
                     /note="10-88-81 : polymorphic base C or T"
BASE COUNT       23 a      6 c      5 g     12 t      1 others
ORIGIN

  Query Match             0.9%;  Score 22.2;  DB 6;  Length 47;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+06;
  Matches   30;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   2107 ATCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGAT 2151
          || |||  || || | || ||| :|||| ||  || | |||| ||
Db     47 ATGTTTGGTTGATTTCAACATAGRTTTTCCAGTTATACTTTTAAT 3


RESULT 33
AX172247/c
LOCUS       AX172247                  33 bp    DNA     linear   PAT 03-JUL-2001
DEFINITION  Sequence 30 from Patent WO0144497.
ACCESSION   AX172247
VERSION     AX172247.1  GI:14597474
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Alessi,D. and Biondi,R.
  TITLE     Protein kinase regulation
  JOURNAL   Patent: WO 0144497-A 30 21-JUN-2001;
            The University of Dundee (GB)
FEATURES             Location/Qualifiers
     source          1. .33
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="PCR primer"
BASE COUNT        8 a     10 c      9 g      6 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 6;  Length 33;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1269 GGCTCCATCTGTACTTGAAAGT 1290
          ||||||||||||||||||||||
Db     33 GGCTCCATCTGTACTTGAAAGT 12


RESULT 34
AR152463
LOCUS       AR152463                  42 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 31 from patent US 6235263.
ACCESSION   AR152463
VERSION     AR152463.1  GI:15119995
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Wong,A.K.C., Bartel,P.L., Teng,D.H.-F. and Tavtigian,S.V.
  TITLE     Carboxy-terminal BRCA1 interacting protein
  JOURNAL   Patent: US 6235263-A 31 22-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="unknown"
BASE COUNT       15 a     11 c     10 g      6 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 6;  Length 42;
  Best Local Similarity   73.7%;  Pred. No. 1.7e+06;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   2034 TTTTCCCAGGCAAAATGGGAAGGCAAAGACAAAAGAAA 2071
          ||||||||| ||  | |||||| ||  |||| | | |
Db      2 TTTTCCCAGTCACGACGGGAAGGAAATCACAAGAAACA 39


RESULT 35
AX164855/c
LOCUS       AX164855                  50 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 50 from Patent WO0138586.
ACCESSION   AX164855
VERSION     AX164855.1  GI:14545684
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0138586-A 50 31-MAY-2001;
            Curagen Corporation (US)
```

```
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    25. .26
                     /note="Nucleotide deleted between bases 25 and 26
                     Accession number cg43056971"
     variation       26
                     /note="single nucleotide polymorphism"
BASE COUNT       16 a      4 c      8 g     22 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 6;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 1.7e+06;
  Matches   31;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2096 AAACAACCTGAATCTTTTTTTTATATAAATATATATTTTTCAAATA 2141
          |||||| ||    |    |||   | |||||  ||||||||||||
Db     48 AAACAATCTTTTACACACTTTGGAAAAAATAGGTATTTTTCAAATA 3


RESULT 36
A51710/c
LOCUS       A51710                    58 bp    DNA     linear   PAT 10-MAR-1997
DEFINITION  Sequence 16 from Patent WO9618744.
ACCESSION   A51710
VERSION     A51710.1  GI:2304514
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Crouzet,J., Scherman,D. and Wils,P.
  TITLE     PURIFICATION OF A TRIPLE HELIX FORMATION WITH AN IMMOBILIZED
            OLIGONUCLEOTIDE
  JOURNAL   Patent: WO 9618744-A 16 20-JUN-1996;
            RHONE POULENC RORER SA (FR)
COMMENT     Other publication AU 4178996 960703
            Other publication FR 2728264 960621.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        2 a      36 c      1 g     19 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 6;  Length 58;
  Best Local Similarity   63.0%;  Pred. No. 1.7e+06;
  Matches   34;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy    123 GGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAGTT 176
          |||  ||   ||||| | ||     |||||  |||  ||||||||  ||   | ||
Db     54 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAATT 1


RESULT 37
AR167589/c
LOCUS       AR167589                  58 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 16 from patent US 6287762.
ACCESSION   AR167589
VERSION     AR167589.1  GI:17903378
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Crouzet,J., Scherman,D. and Wils,P.
  TITLE     Purification of a triple helix formation with an immobilized
            oligonucleotide
  JOURNAL   Patent: US 6287762-A 16 11-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="unknown"
BASE COUNT        2 a      36 c      1 g     19 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 6;  Length 58;
  Best Local Similarity   63.0%;  Pred. No. 1.7e+06;
  Matches   34;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy    123 GGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAGTT 176
          |||  ||   ||||| | ||     |||||  |||  ||||||||  ||   | ||
Db     54 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAATT 1


RESULT 38
AR178299/c
LOCUS       AR178299                  58 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 16 from patent US 6319672.
ACCESSION   AR178299
VERSION     AR178299.1  GI:20219437
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Crouzet,J., Scherman,D., Wils,P., Blanche,F. and Cameron,B.
  TITLE     Purification of a triple helix formation with an immobilized
            oligonucleotide
  JOURNAL   Patent: US 6319672-A 16 20-NOV-2001;
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="unknown"
BASE COUNT        2 a      36 c      1 g     19 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 6;  Length 58;
  Best Local Similarity   63.0%;  Pred. No. 1.7e+06;
  Matches   34;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy    123 GGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAGTT 176
          |||  ||   ||||| | ||     |||||  |||  ||||||||  ||   | ||
Db     54 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAATT 1


RESULT 39
AX323381/c
LOCUS       AX323381                  58 bp    DNA     linear   PAT 07-JAN-2002
DEFINITION  Sequence 16 from Patent WO0192511.
ACCESSION   AX323381
VERSION     AX323381.1  GI:18094143
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Crouzet,J., Scherman,D., Wils,P., Blanche,F. and Cameron,B.
  TITLE     Purification of a triple helix formation with an immobilized
            oligonucleotide
  JOURNAL   Patent: WO 0192511-A 16 06-DEC-2001;
            Aventis Pharma (FR)
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic oligonucleotide"
BASE COUNT        2 a      36 c      1 g     19 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 6;  Length 58;
  Best Local Similarity   63.0%;  Pred. No. 1.7e+06;
  Matches   34;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy    123 GGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAGTT 176
          |||  ||   ||||| | ||     |||||  |||  ||||||||  ||   | ||
Db     54 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAATT 1
```

```
RESULT 40
BD008815
LOCUS       BD008815                  60 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Adhesin binding group antigen.
ACCESSION   BD008815
VERSION     BD008815.1  GI:18637188
KEYWORDS    JP 2001503606-A/3.
SOURCE      Helicobacter pylori.
  ORGANISM  Helicobacter pylori
            Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
            Helicobacter.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Boren,T., Arnqvist,A., Normark,S., Ilver,D. and Hammarstrom,L.
  TITLE     Adhesin binding group antigen
  JOURNAL   Patent: JP 2001503606-A 3 21-MAR-2001;
            THOMAS BOREN
COMMENT     OS   Helicobacter pylori
            PN   JP 2001503606-A/3
            PD   21-MAR-2001
            PF   10-JUN-1997 JP 1998501515
            PR   10-JUN-1996 SE    9602287-6,19-MAR-1997 SE    9701014-4 PI
            THOMAS BOREN,ANNA ARNQVIST,STAFFAN NORMARK,DAG ILVER, PI   LENNART
            HAMMARSTROM
            PC   C07K14/205,A61K39/106,C07K16/12
            CC
            FH   Key             Location/Qualifiers
            FT   source          1. .60
            FT                   /organism='Helicobacter pylori'.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Helicobacter pylori"
                     /db_xref="taxon:210"
BASE COUNT       19 a     13 c     16 g     12 t
ORIGIN

  Query Match             0.9%;  Score 22;  DB 6;  Length 60;
  Best Local Similarity   63.0%;  Pred. No. 1.7e+06;
  Matches   34;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy      49 GACGGCTTTTACCCAGCCCCGGACTTCCGAGACAGGGAAGCTGAGGACATGGCA 102
           ||||||||||||  ||  |   | ||  | |  | | ||||| |   | |||| |
Db       7 GACGGCTTTTACACAAGCGTAGGCTATCAAATCGGTGAAGCCGCTCAAATGGTA 60


RESULT 41
AR117934/c
LOCUS       AR117934                  50 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 7 from patent US 6140470.
ACCESSION   AR117934
VERSION     AR117934.1  GI:14098840
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Garen,A. and Cai,X.
  TITLE     Human monoclonal anti-tumor antibodies
  JOURNAL   Patent: US 6140470-A 7 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="unknown"
BASE COUNT        6 a     11 c     20 g     10 t      3 others
ORIGIN

  Query Match             0.9%;  Score 21.8;  DB 6;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 1.9e+06;
  Matches   29;  Conservative    1;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    1501 GCATCGGCACCACTTCCAATACGACAGCCGAACTCTGGGCCAT 1543
           :|| | ||  ||| ||    ||| | ||||||| ||| |||||
Db      43 SCAMCAGCTGCACCTCGGCCACGTCGGCCGAACCCTGAGCCAT 1


RESULT 42
AR117952/c
LOCUS       AR117952                  50 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 44 from patent US 6140470.
ACCESSION   AR117952
VERSION     AR117952.1  GI:14098858
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Garen,A. and Cai,X.
  TITLE     Human monoclonal anti-tumor antibodies
  JOURNAL   Patent: US 6140470-A 44 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="unknown"
BASE COUNT        6 a     11 c     20 g     10 t      3 others
ORIGIN

  Query Match             0.9%;  Score 21.8;  DB 6;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 1.9e+06;
  Matches   29;  Conservative    1;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    1501 GCATCGGCACCACTTCCAATACGACAGCCGAACTCTGGGCCAT 1543
           :|| | ||  ||| ||    ||| | ||||||| ||| |||||
Db      43 SCAMCAGCTGCACCTCGGCCACGTCGGCCGAACCCTGAGCCAT 1


RESULT 43
AX160479
LOCUS       AX160479                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 3807 from Patent WO0140521.
ACCESSION   AX160479
VERSION     AX160479.1  GI:14541810
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 3807 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (3808 is other entry)
                     Accession number cg43920546"
BASE COUNT       33 a      9 c      5 g      4 t
ORIGIN

  Query Match             0.9%;  Score 21.8;  DB 6;  Length 51;
  Best Local Similarity   70.7%;  Pred. No. 1.9e+06;
  Matches   29;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy     918 TAGAAAGAAAACAATTGACAAAATCCTCAAATGTAAACTCA 958
           |||||| || | ||   | |||| ||||||| | |||  ||
Db      10 TAGAAAAAACAAAAAACAAAAAACCCTCAAAGGAAAAAACA 50


RESULT 44
AX160117/c
LOCUS       AX160117                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 3445 from Patent WO0140521.
ACCESSION   AX160117
VERSION     AX160117.1  GI:14541448
```

```
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 3445 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (3446 is other entry)
                     Accession number cg43271689"
BASE COUNT       13 a      10 c       5 g      23 t
ORIGIN

  Query Match             0.9%;  Score 21.6;  DB 6;  Length 51;
  Best Local Similarity   68.2%;  Pred. No. 2.1e+06;
  Matches   30;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1278 TGTACTTGAAAGTGTGAAAGAAAAGTTTTCCTTTGAACCAAAAA 1321
           |||||   ||| |||||||   ||| |||  ||  ||  |||||
Db     45 TGTACACTAAAATGTGAAATTGAAGGTTTTGTTAAAAAGAAAAA 2


RESULT 45
AX160118/c
LOCUS       AX160118                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 3446 from Patent WO0140521.
ACCESSION   AX160118
VERSION     AX160118.1  GI:14541449
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 3446 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (3445 is other entry)
                     Accession number cg43271689"
BASE COUNT       12 a      10 c       6 g      23 t
ORIGIN

  Query Match             0.9%;  Score 21.6;  DB 6;  Length 51;
  Best Local Similarity   68.2%;  Pred. No. 2.1e+06;
  Matches   30;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1278 TGTACTTGAAAGTGTGAAAGAAAAGTTTTCCTTTGAACCAAAAA 1321
           |||||   ||| |||||||   ||| |||  ||  ||  |||||
Db     45 TGTACACTAAAATGTGAAACTGAAGGTTTTGTTAAAAAGAAAAA 2


RESULT 46
AF079025/c
LOCUS       AF079025                 54 bp    mRNA    linear   PRI 15-SEP-1999
DEFINITION  Homo sapiens clone AV22-1 T-cell receptor alpha chain (TCRA) mRNA,
            partial cds.
ACCESSION   AF079025
VERSION     AF079025.1  GI:5881986
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Fontenot,A.P., Falta,M.T., Newman,L.S. and Kotzin,B.L.
  TITLE     Identification of pathogenic T-cells in patients with
            beryllium-induced lung disease
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 54)
  AUTHORS   Fontenot,A.P., Falta,M.T., Newman,L.S. and Kotzin,B.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-1998) Medicine and Pediatrics, National Jewish
            Medical and Research Center, 1400 Jackson Street, K1020, Denver, CO
            80206, USA
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="AV22-1"
                     /cell_type="CD4+"
                     /note="isolated from peripheral blood and bronchoalveolar
                     lavage of patient 2 with chronic beryllium disease"
     gene            <1. .>54
                     /gene="TCRA"
     CDS             <1. .>54
                     /gene="TCRA"
                     /note="CDR3; 3' end of TCRAV22 to 5' end of AJ8"
                     /codon_start=1
                     /product="T-cell receptor alpha chain"
                     /protein_id="AAD55151.1"
                     /db_xref="GI:5882002"
                     /translation="VYFCALSDNTGFQKLVFG"
BASE COUNT       13 a       9 c      13 g      19 t
ORIGIN

  Query Match             0.9%;  Score 21.6;  DB 9;  Length 54;
  Best Local Similarity   68.2%;  Pred. No. 2.1e+06;
  Matches   30;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1685 TCAAAATGACAGTTTCAGAGAGTCAATGTCATTACATAGAACAC 1728
           || ||||   |||||| || || |  ||| || ||  ||| |||
Db     54 TCCAAATACAAGTTTCTGAAAGCCTGTGTTATCACTCAGAGCAC 11


RESULT 47
AX484751/c
LOCUS       AX484751                 55 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 2051 from Patent WO02053728.
ACCESSION   AX484751
VERSION     AX484751.1  GI:22319035
KEYWORDS    .
SOURCE      Candida albicans.
  ORGANISM  Candida albicans
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 02053728-A 2051 11-JUL-2002;
            Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Candida albicans"
                     /db_xref="taxon:5476"
BASE COUNT       12 a       9 c       9 g      25 t
ORIGIN

  Query Match             0.9%;  Score 21.6;  DB 6;  Length 55;
  Best Local Similarity   68.2%;  Pred. No. 2.1e+06;
  Matches   30;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;
```

```
Qy    913 GAGAATAGAAAGAAAACAATTGACAAAATCCTCAAATGTAAACT 956
          | | | | |||||||||||| |   | |||| |||| | | || |
Db     45 GGGGAGAAAAAGAAAACAAATATGACAATCATCAACTCTTAATT 2


RESULT 48
AX248729
LOCUS       AX248729                  31 bp    DNA     linear   PAT 28-SEP-2001
DEFINITION  Sequence 808 from Patent WO0166800.
ACCESSION   AX248729
VERSION     AX248729.1  GI:15863352
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Cargill,M., Ireland,J.S. and Lander,E.S.
  TITLE     Human single nucleotide polymorphisms
  JOURNAL   Patent: WO 0166800-A 808 13-SEP-2001;
            WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       14 a      5 c      4 g      7 t      1 others
ORIGIN

  Query Match             0.9%;  Score 21.4;  DB 6;  Length 31;
  Best Local Similarity   88.0%;  Pred. No. 2.3e+06;
  Matches   22;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy   1777 TGCAAAAAAAAACTTAAAGCAAAAT 1801
          | ||||||||:||||| ||||||||
Db      6 TCCAAAAAAAARACTTACAGCAAAAT 30


RESULT 49
AR064025
LOCUS       AR064025                  48 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 12 from patent US 5846774.
ACCESSION   AR064025
VERSION     AR064025.1  GI:5993333
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Xia,Y.
  TITLE     Chlorella virus promoters
  JOURNAL   Patent: US 5846774-A 12 08-DEC-1998;
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="unknown"
BASE COUNT       17 a      7 c      9 g     15 t
ORIGIN

  Query Match             0.9%;  Score 21.4;  DB 6;  Length 48;
  Best Local Similarity   66.0%;  Pred. No. 2.4e+06;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   2233 CTTCTCTTGGTGAAATAATAAAATGCAAATGAATCATTGTTAACACA 2279
          ||| || || ||| || |||||   |||||||  |  |||    |||
Db      2 CTTATCGTGTTGATATGATAAATGACAAATGATACGCTGTATCAACA 48


RESULT 50
AR178317
LOCUS       AR178317                  48 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 34 from patent US 6319672.
ACCESSION   AR178317
VERSION     AR178317.1  GI:20219455
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Crouzet,J., Scherman,D., Wils,P., Blanche,F. and Cameron,B.
  TITLE     Purification of a triple helix formation with an immobilized
            oligonucleotide
  JOURNAL   Patent: US 6319672-A 34 20-NOV-2001;
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="unknown"
BASE COUNT       16 a      0 c      32 g      0 t
ORIGIN

  Query Match             0.9%;  Score 21.4;  DB 6;  Length 48;
  Best Local Similarity   66.0%;  Pred. No. 2.4e+06;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    123 GGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGG 169
          |||  ||   ||||| | ||     ||||| |||  |||||||| ||
Db      1 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 47


RESULT 51
AX323399
LOCUS       AX323399                  48 bp    DNA     linear   PAT 07-JAN-2002
DEFINITION  Sequence 34 from Patent WO0192511.
ACCESSION   AX323399
VERSION     AX323399.1  GI:18094161
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Crouzet,J., Scherman,D., Wils,P., Blanche,F. and Cameron,B.
  TITLE     Purification of a triple helix formation with an immobilized
            oligonucleotide
  JOURNAL   Patent: WO 0192511-A 34 06-DEC-2001;
            Aventis Pharma (FR)
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic oligonucleotide"
BASE COUNT       16 a      0 c      32 g      0 t
ORIGIN

  Query Match             0.9%;  Score 21.4;  DB 6;  Length 48;
  Best Local Similarity   66.0%;  Pred. No. 2.4e+06;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    123 GGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGG 169
          |||  ||   ||||| | ||     ||||| |||  |||||||| ||
Db      1 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 47


RESULT 52
AX162478/c
LOCUS       AX162478                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 5806 from Patent WO0140521.
ACCESSION   AX162478
VERSION     AX162478.1  GI:14543809
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 5806 07-JUN-2001;
            Curagen Corporation (US)
```

```
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (5805 is other entry)
                     Accession number cg44035239"
BASE COUNT       18 a      2 c      5 g     26 t
ORIGIN

  Query Match             0.9%;  Score 21.4;  DB 6;  Length 51;
  Best Local Similarity   66.0%;  Pred. No. 2.4e+06;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   2097 AACAACCTGAATCTTTTTTTTATATAAATATATATTTTTCAAATAGA 2143
          |||||  |  || | | || | |||| |||||||| |  |  ||| |
Db     48 AACAAAATATATATATATTCTCTATATATATATATATAGCTGATAAA 2

RESULT 53
I29928
LOCUS       I29928                    55 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 41 from patent US 5578468.
ACCESSION   I29928
VERSION     I29928.1  GI:1820719
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Pickup,D.J., Patel,D. and Antczak,J.B.
  TITLE     Site-specific RNA cleavage
  JOURNAL   Patent: US 5578468-A 41 26-NOV-1996;
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="unknown"
BASE COUNT       47 a      2 c      1 g      5 t
ORIGIN

  Query Match             0.9%;  Score 21.4;  DB 6;  Length 55;
  Best Local Similarity   66.0%;  Pred. No. 2.4e+06;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   1754 GATTTTAAAAAATCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAA 1800
          ||||||| ||||  ||  ||       |||||||||   |||  ||||
Db      8 GATTTTATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 54

RESULT 54
I66359
LOCUS       I66359                    60 bp    DNA     linear   PAT 28-DEC-1997
DEFINITION  Sequence 18 from patent US 5670330.
ACCESSION   I66359
VERSION     I66359.1  GI:2724336
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Sonenberg,N., Katze,M.G., Roy,S., Koromilas,A.E. and Barber,G.H.
  TITLE     Anti-tumor agent assay using PKR
  JOURNAL   Patent: US 5670330-A 18 23-SEP-1997;
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="unknown"
BASE COUNT       26 a      6 c     10 g     18 t
ORIGIN

  Query Match             0.9%;  Score 21.4;  DB 6;  Length 60;
  Best Local Similarity   61.8%;  Pred. No. 2.4e+06;
  Matches   34;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy    644 TGGGGCATTTACATCAAAAGGGGATCATCTACAGAGACCTGAAGCCGGAGAATAT 698
          |||   || |||||  |||     | ||  | ||||| || |||||    |||||
Db      5 TGGATTATATACATTCAAAAAAATTAATTCATAGAGATCTTAAGCCAAGTAATAT 59

RESULT 55
AX160974/c
LOCUS       AX160974                  50 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 4302 from Patent WO0140521.
ACCESSION   AX160974
VERSION     AX160974.1  GI:14542305
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 4302 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    25. .26
                     /note="Nucleotide deleted between bases 25 and 26
                     Accession number cg43944408"
     misc_feature    26
                     /note="2 of 2 allelic variants (4301 is other entry)"
BASE COUNT       12 a     13 c      7 g     18 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 6;  Length 50;
  Best Local Similarity   69.0%;  Pred. No. 2.6e+06;
  Matches   29;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    895 GCACCCCCATTCACTGGGGAGAATAGAAAGAAAACAATTGAC 936
          |||  || ||| ||||||| ||| | ||||  || |  || |
Db     46 GCATACCAATTTACTGGGGGGAAAAAAAAGTTAAGAGATGCC 5

RESULT 56
AX157225/c
LOCUS       AX157225                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 553 from Patent WO0140521.
ACCESSION   AX157225
VERSION     AX157225.1  GI:14538556
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 553 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (554 is other entry)
                     Accession number cg44938869"
BASE COUNT       25 a      2 c      9 g     15 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 6;  Length 51;
  Best Local Similarity   64.0%;  Pred. No. 2.6e+06;
  Matches   32;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;
```

```
Qy   2096 AAACAACCTGAATCTTTTTTTTATATAAATATATATTTTTCAAATAGATT 2145
          ||||||||   || ||||||||  | |  ||   | || ||  |  | |||
Db     51 AAACAACCCTGATTTTTTTTTTTTTTGCATCATTTTTATTACACCAAATT 2

RESULT 57
AX160455/c
LOCUS       AX160455                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 3783 from Patent WO0140521.
ACCESSION   AX160455
VERSION     AX160455.1  GI:14541786
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 3783 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (3784 is other entry)
                     Accession number cg43920465"
BASE COUNT       24 a      6 c      8 g     13 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 6;  Length 51;
  Best Local Similarity   64.0%;  Pred. No. 2.6e+06;
  Matches   32;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy   2102 CCTGAATCTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGAT 2151
          | | | | || |  ||||  | || ||  ||| | |||||| | ||| ||
Db     51 CTTTATTGTTCTAATTATCCAGATGTACCTTTGTAAAATAGCTCTTTTAT 2

RESULT 58
AR192747
LOCUS       AR192747                  54 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 8235 from patent US 6346398.
ACCESSION   AR192747
VERSION     AR192747.1  GI:20238712
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Pavco,P., McSwiggen,J., Stinchcomb,D. and Escobedo,J.
  TITLE     Method and reagent for the treatment of diseases or conditions
            related to levels of vascular endothelial growth factor receptor
  JOURNAL   Patent: US 6346398-A 8235 12-FEB-2002;
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="unknown"
BASE COUNT       21 a      9 c     14 g     10 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 6;  Length 54;
  Best Local Similarity   64.0%;  Pred. No. 2.7e+06;
  Matches   32;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy   2045 AAAATGGGAAGGCAAAGACAAAAGAAACTTACCAATTGATGTTTTACGTG 2094
          ||  |||  | ||||||||  | ||||||  ||    || |   ||||  ||
Db      2 AATGTGGAGAAGCAAAGACCAGAGAAACACACGTTGTGGTACATTACCTG 51

RESULT 59
AF254565/c
LOCUS       AF254565                  54 bp    DNA     linear   PRI 30-OCT-2000
DEFINITION  Homo sapiens ATP7B (ATP7B) gene, exon 8 and  partial cds.
ACCESSION   AF254565
VERSION     AF254565.1  GI:11037272
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Wang,N., Wu,Z.Y., MuRong,S.X., Lin,M.T. and Fang,L.
  TITLE     Hot point mutations of Wilson disease gene in Chinese with DNA
            sequencing
  JOURNAL   Chung-Hua Shen Ching Ko Tsa Chih 31, 20-23 (1998)
REFERENCE   2  (bases 1 to 54)
  AUTHORS   Wu,Z.Y., Wang,N. and MuRong,S.X.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) Department of Neurology, First Affiliated
            Hospital of Fujian Medical University, Chazhong Road 20, Fuzhou,
            Fujian 350005, P.R. China
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /map="13q14.3"
     gene            <1. .>54
                     /gene="ATP7B"
     mRNA            <1. .>54
                     /gene="ATP7B"
                     /product="ATP7B"
     CDS             <1. .>54
                     /gene="ATP7B"
                     /codon_start=1
                     /product="ATP7B"
                     /protein_id="AAG27542.1"
                     /db_xref="GI:11037273"
                     /translation="DTPPMLFVFIALGLWLEH"
     exon            <1. .>54
                     /gene="ATP7B"
                     /number=8
     variation       41
                     /gene="ATP7B"
                     /note="compared to sequence with GenBank Accession Number
                     U11700"
                     /replace="g"
BASE COUNT       7 a      18 c     15 g     14 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 9;  Length 54;
  Best Local Similarity   64.0%;  Pred. No. 2.7e+06;
  Matches   32;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy    122 TGGACCAGCCAGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGT 171
          ||  ||||||| ||| |  ||||  |||  |  | | | | | |||| ||
Db     53 TGTTCCAGCCACAGGCCCAGGGCAATGAACACAAACAGCATGGGGGGCGT 4

RESULT 60
A43625/c
LOCUS       A43625                    55 bp    DNA     linear   PAT 06-MAR-1997
DEFINITION  Sequence 15 from Patent WO9507987.
ACCESSION   A43625
VERSION     A43625.1  GI:2298827
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Bublot,M., De,W.C., Colau,D. and Roux-Salembien,P.
  TITLE     NOVEL PROTEINS/POLYPEPTIDES AND COTRANSFECTION PLASMIDS AND LIVE
            RECOMBINANT CARRIERS THEREFOR
  JOURNAL   Patent: WO 9507987-A 15 23-MAR-1995;
```

```
            SOLVAY (BE)
COMMENT     Other publication NO 961086 960509
            Other publication ZA 9406887 950627
            Other publication AU 7615894 950403
            Other publication GB 2282601 950412.
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       19 a      10 c       8 g      18 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 6;  Length 55;
  Best Local Similarity   64.0%;  Pred. No. 2.7e+06;
  Matches   32;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy   1273 CCATCTGTACTTGAAAGTGTGAAAGAAAAGTTTTCCTTTGAACCAAAAAT 1322
           | |  ||||||| |  ||| | |      |||||  ||  ||||||||||
Db     50 CAAAATGTACTTCATGGTGGGGATCCTTTGTTTTAGTTCAAACCAAAAAT 1

RESULT 61
A80902
LOCUS       A80902                    56 bp    DNA     linear   PAT 21-JAN-2000
DEFINITION  Sequence 13 from Patent EP0919631.
ACCESSION   A80902
VERSION     A80902.1  GI:6731507
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   Goetz,F.P. and Pohlner,J.D.
  TITLE     Process for the identification of nucleic acids
  JOURNAL   Patent: EP 0919631-A 13 02-JUN-1999;
            EVOTEC BIOSYSTEMS AG (DE)
FEATURES             Location/Qualifiers
     source          1. .56
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       23 a      11 c      12 g      10 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 6;  Length 56;
  Best Local Similarity   69.0%;  Pred. No. 2.7e+06;
  Matches   29;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    221 GCATGGAACATTGTGAGAAATTTGAAATCTCAGAAACTAGTG 262
          || |   ||| | | |||||| |||||| ||  || | ||||
Db      9 GCTTACCACAATCTAAGAAATCTGAAATATCTCAAGCAAGTG 50

RESULT 62
A97341
LOCUS       A97341                    56 bp    DNA     linear   PAT 26-JAN-2000
DEFINITION  Sequence 13 from Patent WO9916900.
ACCESSION   A97341
VERSION     A97341.1  GI:6780700
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   Goetz,F. and Strauss,A.
  TITLE     METHOD FOR IDENTIFYING A NUCLEIC ACID
  JOURNAL   Patent: WO 9916900-A 13 08-APR-1999;
            GOETZ FRIEDRICH (DE); STRAUSS ANDREAS (DE)
FEATURES             Location/Qualifiers
     source          1. .56
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       23 a      11 c      12 g      10 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 6;  Length 56;
  Best Local Similarity   69.0%;  Pred. No. 2.7e+06;
  Matches   29;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    221 GCATGGAACATTGTGAGAAATTTGAAATCTCAGAAACTAGTG 262
          || |   ||| | | |||||| |||||| ||  || | ||||
Db      9 GCTTACCACAATCTAAGAAATCTGAAATATCTCAAGCAAGTG 50

RESULT 63
A97358
LOCUS       A97358                    56 bp    DNA     linear   PAT 26-JAN-2000
DEFINITION  Sequence 13 from Patent WO9916894.
ACCESSION   A97358
VERSION     A97358.1  GI:6780716
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   Goetz,F. and Strauss,A.
  TITLE     METHOD FOR DETERMINING ACTIVE AGENTS
  JOURNAL   Patent: WO 9916894-A 13 08-APR-1999;
            GOETZ FRIEDRICH (DE); STRAUSS ANDREAS (DE)
FEATURES             Location/Qualifiers
     source          1. .56
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       23 a      11 c      12 g      10 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 6;  Length 56;
  Best Local Similarity   69.0%;  Pred. No. 2.7e+06;
  Matches   29;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    221 GCATGGAACATTGTGAGAAATTTGAAATCTCAGAAACTAGTG 262
          || |   ||| | | |||||| |||||| ||  || | ||||
Db      9 GCTTACCACAATCTAAGAAATCTGAAATATCTCAAGCAAGTG 50

RESULT 64
AR183071
LOCUS       AR183071                  56 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 13 from patent US 6339174.
ACCESSION   AR183071
VERSION     AR183071.1  GI:20226278
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   Bogdanovic,S.
  TITLE     Method for preparing aldehydes by hydroformylation
  JOURNAL   Patent: US 6339174-A 13 15-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .56
                     /organism="unknown"
BASE COUNT       23 a      11 c      12 g      10 t
ORIGIN

  Query Match             0.9%;  Score 21.2;  DB 6;  Length 56;
  Best Local Similarity   69.0%;  Pred. No. 2.7e+06;
  Matches   29;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    221 GCATGGAACATTGTGAGAAATTTGAAATCTCAGAAACTAGTG 262
          || |   ||| | | |||||| |||||| ||  || | ||||
Db      9 GCTTACCACAATCTAAGAAATCTGAAATATCTCAAGCAAGTG 50

RESULT 65
E07494
LOCUS       E07494                    38 bp    DNA     linear   PAT 29-SEP-1997
```

```
DEFINITION  Synthetic DNA for probe.
ACCESSION   E07494
VERSION     E07494.1  GI:2175632
KEYWORDS    JP 1994133799-A/3.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Yamanishi,K., Yamamoto,T. and Mori,H.
  TITLE     ANALYSIS OF HUMAN HERPES VIRUS 6 TYPE @(3754/24)HHV-6) DNA AND
            DISCRIMINATION OF SUB-TYPE
  JOURNAL   Patent: JP 1994133799-A 3 17-MAY-1994;
            INTERNATL REAGENTS CORP
COMMENT     OS   None
            OC   Artificial sequences.
            PN   JP 1994133799-A/3
            PD   17-MAY-1994
            PF   27-OCT-1992 JP 1992311416
            PI   YAMANISHI KOICHI, YAMAMOTO TAKESHI, MORI HIROYUKI PC
            C12Q1/68,C12Q1/68,C12N15/11,C12N15/38;
            CC   strandedness: Single;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .38
            FT                   /organism='Artificial sequences'.
FEATURES              Location/Qualifiers
     source           1. .38
                      /organism="unidentified"
                      /db_xref="taxon:32644"
BASE COUNT       16 a      5 c      5 g     12 t
ORIGIN

  Query Match             0.9%;  Score 21;  DB 6;  Length 38;
  Best Local Similarity   73.0%;  Pred. No. 2.9e+06;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1749 ACGTGGATTTTAAAAAATCAATCAATGGTGCAAAAAA 1785
          | ||| ||| |||||| | ||| | ||  || |||||
Db      1 ATGTGAATTGTAAAAATTTAATTACTGCCGCCAAAAA 37


RESULT 66
AX299737/c
LOCUS       AX299737                 40 bp    DNA     linear   PAT 26-NOV-2001
DEFINITION  Sequence 17 from Patent WO0175163.
ACCESSION   AX299737
VERSION     AX299737.1  GI:17129277
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Landers,J.E.
  TITLE     High throughput methods for haplotyping
  JOURNAL   Patent: WO 0175163-A 17 11-OCT-2001;
            Polygenyx, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .40
                      /organism="synthetic construct"
                      /db_xref="taxon:32630"
                      /note="Synthetic Oligonucleotide"
     misc_feature     1
                      /note="amino group attached"
BASE COUNT        7 a      4 c      3 g     26 t
ORIGIN

  Query Match             0.9%;  Score 21;  DB 6;  Length 40;
  Best Local Similarity   73.0%;  Pred. No. 2.9e+06;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1766 TCAATCAATGGTGCAAAAAAAAACTTAAAGCAAAATA 1802
          ||| ||| ||  |||||||||||   |||  |||| |
Db     37 TCATTCATTGTGGCAAAAAAAAAAAAAAAAAAAAAAA 1


RESULT 67
AX160105
LOCUS       AX160105                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 3433 from Patent WO0140521.
ACCESSION   AX160105
VERSION     AX160105.1  GI:14541436
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 3433 07-JUN-2001;
            Curagen Corporation (US)
FEATURES              Location/Qualifiers
     source           1. .51
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
     misc_feature     26
                      /note="1 of 2 allelic variants (3434 is other entry)
                      Accession number cg43268348"
BASE COUNT       25 a      6 c      6 g     14 t
ORIGIN

  Query Match             0.9%;  Score 21;  DB 6;  Length 51;
  Best Local Similarity   66.7%;  Pred. No. 2.9e+06;
  Matches   30;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   1733 GACACAGGAAAAATAAACGTGGATTTTAAAAAATCAATCAATGGT 1777
          || |||| ||||| |||  | | ||| ||| |   ||||  | ||
Db      6 GAAACAGTAAAAAAAAAAATAGGTTTGAAACACATAATCCTTCGT 50


RESULT 68
MUSBMP24H
LOCUS       MUSBMP24H                56 bp    DNA     linear   ROD 29-MAY-2002
DEFINITION  Mouse gene for BMP-2/4 type I receptor, exon5 (partial) and intron5
            (partial).
ACCESSION   D45009
VERSION     D45009.1  GI:624892
KEYWORDS    BMP-2/4 type I receptor.
SOURCE      Mus musculus (sub_species:domesticus, strain:129) adult male liver
            DNA, clone_lib:lambda DASH-II.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   Mishina,Y., Suzuki,A., Debra,G.J., Copeland,N.G., Jenkins,N.A.,
            Ueno,N. and Behringer,R.R.
  TITLE     Genomic organization and chromosomal location of the mouse type I
            BMP-2/4 receptor
  JOURNAL   Biochem. Biophys. Res. Commun. (1995) In press
REFERENCE   2  (bases 1 to 56)
  AUTHORS   Behringer,R.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-DEC-1994) Richard R. Behringer, The University of
            Texas M.D.Anderson Cancer Center, Molecular Genetics; 1515
            Holcombe, Houston, Texas 77030, U.S.A.
            (E-mail:sa11112@odin.mda.uth.tmc.edu, Tel:713-794-4631,
            Fax:713-794-4394)
FEATURES              Location/Qualifiers
     source           1. .56
                      /organism="Mus musculus"
                      /strain="129"
                      /sub_species="domesticus"
```

```
                     /db_xref="taxon:10090"
                     /chromosome="14"
                     /sex="male"
                     /tissue_type="liver"
                     /clone_lib="lambda DASH-II"
                     /dev_stage="adult"
     exon            <1. .10
                     /product="BMP-2/4 type I receptor"
                     /note="the length of exon5 is 100 bp"
                     /number=5
     intron          11. .>56
                     /note="the lentgh of intron5 is 4.8 kb"
                     /number=5
BASE COUNT       19 a      5 c      4 g     28 t
ORIGIN

  Query Match             0.9%;  Score 21;  DB 10;  Length 56;
  Best Local Similarity   73.0%;  Pred. No. 3e+06;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   2093 TGCAAACAACCTGAATCTTTTTTTTATATAAATATAT 2129
          ||| |  |    || |||| |||||||| ||||||||
Db      3 TGCTATAAGTAAGATTCTTATTTTTATACAAATATAT 39


RESULT 69
AR209950
LOCUS       AR209950                  59 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 4 from patent US 6387631.
ACCESSION   AR209950
VERSION     AR209950.1  GI:21512053
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Arnold,L.J., Sawan,S.P. and Lee,P.H.
  TITLE     Polymer coated surfaces for microarray applications
  JOURNAL   Patent: US 6387631-A 4 14-MAY-2002;
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="unknown"
BASE COUNT       25 a      4 c     12 g     18 t
ORIGIN

  Query Match             0.9%;  Score 21;  DB 6;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 3e+06;
  Matches   30;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   1078 TTCTTTAGACACATTAACTGGGAAGAACTTCTGGCTCGAAAGGTG 1122
          ||| ||||  |||||||  |   | ||||  ||  | |||| |||
Db      5 TTCATTAGTGACATTAAAAGAAGAAAACTGATGTTTTGAAATGTG 49


RESULT 70
AX253387
LOCUS       AX253387                  59 bp    DNA     linear   PAT 10-OCT-2001
DEFINITION  Sequence 17 from Patent WO0171039.
ACCESSION   AX253387
VERSION     AX253387.1  GI:16073921
KEYWORDS    .
SOURCE      Candida albicans.
  ORGANISM  Candida albicans
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Reynolds,M.A., Ruvolo,M. and Arnold,L.J.
  TITLE     Combined polynucleotide sequences as discrete assay endpoints
  JOURNAL   Patent: WO 0171039-A 17 27-SEP-2001;
            Incyte Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Candida albicans"
                     /db_xref="taxon:5476"
                     /note="Incyte ID No: F23S805"
BASE COUNT       25 a      4 c     12 g     18 t
ORIGIN

  Query Match             0.9%;  Score 21;  DB 6;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 3e+06;
  Matches   30;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   1078 TTCTTTAGACACATTAACTGGGAAGAACTTCTGGCTCGAAAGGTG 1122
          ||| ||||  |||||||  |   | ||||  ||  | |||| |||
Db      5 TTCATTAGTGACATTAAAAGAAGAAAACTGATGTTTTGAAATGTG 49


RESULT 71
AX254767
LOCUS       AX254767                  59 bp    DNA     linear   PAT 10-OCT-2001
DEFINITION  Sequence 4 from Patent WO0170641.
ACCESSION   AX254767
VERSION     AX254767.1  GI:16074429
KEYWORDS    .
SOURCE      Candida albicans.
  ORGANISM  Candida albicans
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Arnold,L.J., Sawan,S.P. and Lee,P.H.
  TITLE     Polymer coated surfaces for microarray applications
  JOURNAL   Patent: WO 0170641-A 4 27-SEP-2001;
            Incyte Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Candida albicans"
                     /db_xref="taxon:5476"
                     /note="Incyte ID No: F23S805"
BASE COUNT       25 a      4 c     12 g     18 t
ORIGIN

  Query Match             0.9%;  Score 21;  DB 6;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 3e+06;
  Matches   30;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   1078 TTCTTTAGACACATTAACTGGGAAGAACTTCTGGCTCGAAAGGTG 1122
          ||| ||||  |||||||  |   | ||||  ||  | |||| |||
Db      5 TTCATTAGTGACATTAAAAGAAGAAAACTGATGTTTTGAAATGTG 49


RESULT 72
AR073297
LOCUS       AR073297                  60 bp    DNA     linear   PAT 28-AUG-2000
DEFINITION  Sequence 2 from patent US 5948894.
ACCESSION   AR073297
VERSION     AR073297.1  GI:10000060
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Berry,M.J., Davis,P.J., Verhoeyen,M.E. and De Winter,R.F.J.
  TITLE     Immunoadsorbent reagents
  JOURNAL   Patent: US 5948894-A 2 07-SEP-1999;
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="unknown"
BASE COUNT       12 a     16 c     13 g     13 t      6 others
ORIGIN

  Query Match             0.9%;  Score 21;  DB 6;  Length 60;
  Best Local Similarity   62.8%;  Pred. No. 3e+06;
  Matches   27;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy     58 TACCCAGCCCCGGACTTCCGAGACAGGGAAGCTGAGGACATGG 100
          |||||    ||||| ||||       |||  ||||||| | ||
```

```
Db     17 TACCCCTTACCGGAATTCCCNNNNNNGGATCCTGAGGAGACGG 59


RESULT 73
YSCMTP191/c
LOCUS       YSCMTP191                 60 bp    DNA     linear   PLN 04-AUG-1993
DEFINITION  Yeast (S.cerevisiae) mitochondrial petite mutant excision seq 19,
            left end.
ACCESSION   J01506
VERSION     J01506.1  GI:343885
KEYWORDS    AT-rich region; GC rich region.
SEGMENT     1 of 2
SOURCE      Yeast (Saccharomyces cerevisiae) mitochondrial DNA.
  ORGANISM  Mitochondrion Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   de Zamaroczy,M., Faugeron-Fonty,G. and Bernardi,G.
  TITLE     Excision sequences in the mitochondrial genome of yeast
  JOURNAL   Gene 21 (3), 193-202 (1983)
  MEDLINE   83210931
   PUBMED   6343188
COMMENT     Additional sequences reported in [1], but sequenced in earlier
            papers, appear in separate entries. Excision repeat corresponds to
            bases 11 to 36.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Saccharomyces cerevisiae"
                     /organelle="mitochondrion"
                     /db_xref="taxon:4932"
BASE COUNT       32 a      2 c      0 g     26 t
ORIGIN

  Query Match             0.9%;  Score 21;  DB 8;  Length 60;
  Best Local Similarity   82.8%;  Pred. No. 3e+06;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy   2111 TTTTTTTATATAAATATATATTTTTCAAA 2139
           | ||| |||||| ||||||||||||  ||
Db     37 TATTTATATATATATATATATTTTTATAA 9


RESULT 74
AF177252
LOCUS       AF177252                  57 bp    RNA     linear   INV 21-FEB-2000
DEFINITION  Bodo saltans clone pBG764 mitochondrial putative gRNA.
ACCESSION   AF177252
VERSION     AF177252.1  GI:7012655
KEYWORDS    .
SOURCE      Bodo saltans.
  ORGANISM  Mitochondrion Bodo saltans
            Eukaryota; Euglenozoa; Kinetoplastida; Bodonidae; Bodo.
REFERENCE   1  (bases 1 to 57)
  AUTHORS   Blom,D., de Haan,A., van den Burg,J., van den Berg,M., Sloof,P.,
            Jirku,M., Lukes,J. and Benne,R.
  TITLE     Mitochondrial minicircles in the free-living bodonid Bodo saltans
            contain two gRNA gene cassettes and are not found in large networks
  JOURNAL   RNA 6 (1), 121-135 (2000)
  MEDLINE   20132239
   PUBMED   10668805
REFERENCE   2  (bases 1 to 57)
  AUTHORS   Blom,D., De Haan,A., Sloof,P., Jirku,M., Lukes,J. and Benne,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-AUG-1999) Department of Biochemistry, Academic
            Medical Center, Meibergdreef 15, Amsterdam 1105 AZ, The Netherlands
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="Bodo saltans"
                     /organelle="mitochondrion"
                     /db_xref="taxon:75058"
                     /chromosome="minicircle"
                     /clone="pBG764"
     misc_RNA        1. .57
                     /note="gRNA"
                     /evidence=not_experimental
BASE COUNT       23 a      10 c      4 g     20 t
ORIGIN

  Query Match             0.9%;  Score 20.8;  DB 3;  Length 57;
  Best Local Similarity   60.7%;  Pred. No. 3.3e+06;
  Matches   34;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy   2062 ACAAAAGAAACTTACCAATTGATGTTTTACGTGCAAACAACCTGAATCTTTTTTTT 2117
           ||||||  |||      |||||  | |  |  | ||||||  |   | ||||||||
Db      2 ACAAAAACAACAACAGCATTGAAGCTATCAACGAAAACAATTTTTTTTTTTTTTTT 57


RESULT 75
A48423/c
LOCUS       A48423                    28 bp    DNA     linear   PAT 07-MAR-1997
DEFINITION  Sequence 46 from Patent WO9603501.
ACCESSION   A48423
VERSION     A48423.1  GI:2302214
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Chaudhuri,B., Stephan,C. and Fuerst,P.
  TITLE     DUAL HYBRID SYSTEM
  JOURNAL   Patent: WO 9603501-A 46 08-FEB-1996;
            CIBA GEIGY AG (CH)
COMMENT     Other publication AU 2983295 960222.
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     mat_peptide     complement(9. .28)
                     /note="5'END ANTI-SENSE STRAND OF CODING SEQUENCE OF RAT
                     P70S6 KINASE GENE"
                     /function="PCR FRAGMENT"
BASE COUNT       11 a       5 c       6 g       6 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 28;
  Best Local Similarity   85.2%;  Pred. No. 3.5e+06;
  Matches   23;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy   1376 AATTTTCTCCTGGGGATTTCTGGGGAA 1402
           |||| ||||||||||||||||   |||
Db     28 AATTCTCTCCTGGGGATTTCTCTAGAA 2


RESULT 76
A48424/c
LOCUS       A48424                    28 bp    DNA     linear   PAT 07-MAR-1997
DEFINITION  Sequence 47 from Patent WO9603501.
ACCESSION   A48424
VERSION     A48424.1  GI:2302215
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Chaudhuri,B., Stephan,C. and Fuerst,P.
  TITLE     DUAL HYBRID SYSTEM
  JOURNAL   Patent: WO 9603501-A 47 08-FEB-1996;
            CIBA GEIGY AG (CH)
COMMENT     Other publication AU 2983295 960222.
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     mat_peptide     complement(9. .28)
                     /note="5'END ANTI-SENSE STRAND OF CODING SEQUENCE OF
                     TRUNCATED P70S6K DELTA C GENE"
```

```
                    /function="PCR FRAGMENT"
BASE COUNT        9 a      5 c      5 g      9 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 28;
  Best Local Similarity   85.2%;  Pred. No. 3.5e+06;
  Matches   23;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy   1586 ACCTGCGTATGAATCTATGACAGAGCA 1612
          |||||||||||||||||||  |    ||
Db     28 ACCTGCGTATGAATCTATGAGAATTCA 2


RESULT 77
A51711
LOCUS       A51711                   50 bp    DNA     linear   PAT 10-MAR-1997
DEFINITION  Sequence 17 from Patent WO9618744.
ACCESSION   A51711
VERSION     A51711.1  GI:2304515
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Crouzet,J., Scherman,D. and Wils,P.
  TITLE     PURIFICATION OF A TRIPLE HELIX FORMATION WITH AN IMMOBILIZED
            OLIGONUCLEOTIDE
  JOURNAL   Patent: WO 9618744-A 17 20-JUN-1996;
            RHONE POULENC RORER SA (FR)
COMMENT     Other publication AU 4178996 960703
            Other publication FR 2728264 960621.
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       17 a      0 c     33 g      0 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 3.6e+06;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    132 AGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAG 174
          |||||| | ||     ||||| |||  |||||||| ||   ||
Db      3 AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 45


RESULT 78
AR167590
LOCUS       AR167590                 50 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 17 from patent US 6287762.
ACCESSION   AR167590
VERSION     AR167590.1  GI:17903379
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Crouzet,J., Scherman,D. and Wils,P.
  TITLE     Purification of a triple helix formation with an immobilized
            oligonucleotide
  JOURNAL   Patent: US 6287762-A 17 11-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="unknown"
BASE COUNT       17 a      0 c     33 g      0 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 3.6e+06;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    132 AGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAG 174
          |||||| | ||     ||||| |||  |||||||| ||   ||
Db      3 AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 45


RESULT 79
AR178300
LOCUS       AR178300                 50 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 17 from patent US 6319672.
ACCESSION   AR178300
VERSION     AR178300.1  GI:20219438
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Crouzet,J., Scherman,D., Wils,P., Blanche,F. and Cameron,B.
  TITLE     Purification of a triple helix formation with an immobilized
            oligonucleotide
  JOURNAL   Patent: US 6319672-A 17 20-NOV-2001;
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="unknown"
BASE COUNT       17 a      0 c     33 g      0 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 3.6e+06;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    132 AGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAG 174
          |||||| | ||     ||||| |||  |||||||| ||   ||
Db      3 AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 45


RESULT 80
AX165034
LOCUS       AX165034                 50 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 229 from Patent WO0138586.
ACCESSION   AX165034
VERSION     AX165034.1  GI:14545863
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0138586-A 229 31-MAY-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    25. .26
                     /note="Nucleotide deleted between bases 25 and 26
                     Accession number cg42907867"
     variation       26
                     /note="single nucleotide polymorphism"
BASE COUNT        9 a      8 c     21 g     12 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 50;
  Best Local Similarity   74.3%;  Pred. No. 3.6e+06;
  Matches   26;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy   2301 TGGGGTGGACCTGGGGTTTATTTTCAGTAACCCAG 2335
          ||||  ||| |||| |||  ||||| | || ||||
Db     15 TGGGAGGGATCTGGCGTTGGTTTTCTGAAAGCCAG 49


RESULT 81
```

```
AX323382
LOCUS       AX323382                  50 bp    DNA     linear   PAT 07-JAN-2002
DEFINITION  Sequence 17 from Patent WO0192511.
ACCESSION   AX323382
VERSION     AX323382.1  GI:18094144
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Crouzet,J., Scherman,D., Wils,P., Blanche,F. and Cameron,B.
  TITLE     Purification of a triple helix formation with an immobilized
            oligonucleotide
  JOURNAL   Patent: WO 0192511-A 17 06-DEC-2001;
            Aventis Pharma (FR)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic oligonucleotide"
BASE COUNT       17 a      0 c     33 g      0 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 3.6e+06;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy     132 AGAGGACGCGGGCTCTGAGGATGAGCTGGAGGAGGGGGGTCAG 174
           |||||| | ||     ||||| ||| |||||||| ||   ||
Db       3 AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 45


RESULT 82
AR068824
LOCUS       AR068824                  51 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 24 from patent US 5854051.
ACCESSION   AR068824
VERSION     AR068824.1  GI:6001031
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Chandrashekar,R. and Tsuji,N.
  TITLE     Parasitic helminth asparaginase proteins, nucleic acid molecules,
            and uses thereof
  JOURNAL   Patent: US 5854051-A 24 29-DEC-1998;
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="unknown"
BASE COUNT       12 a      12 c      5 g      22 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 51;
  Best Local Similarity   67.4%;  Pred. No. 3.6e+06;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    1848 CATCTTTCTCAACCTTATCAAGGATTTTCATGTTGATGACTCG 1890
           | |||| || ||| || | |   |||||||  | |||||| |
Db       9 CTTCTTACTGAACTTTTTTCATCTTTTTCATTCTAATGACTAG 51


RESULT 83
AR122557
LOCUS       AR122557                  51 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 24 from patent US 6165735.
ACCESSION   AR122557
VERSION     AR122557.1  GI:14106874
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Chandrashekar,R. and Tsuji,N.
  TITLE     Parasitic helminth asparaginase proteins, nucleic acid molecules,
            and uses thereof
  JOURNAL   Patent: US 6165735-A 24 26-DEC-2000;
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="unknown"
BASE COUNT       12 a      12 c      5 g      22 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 51;
  Best Local Similarity   67.4%;  Pred. No. 3.6e+06;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    1848 CATCTTTCTCAACCTTATCAAGGATTTTCATGTTGATGACTCG 1890
           | |||| || ||| || | |   |||||||  | |||||| |
Db       9 CTTCTTACTGAACTTTTTTCATCTTTTTCATTCTAATGACTAG 51


RESULT 84
AX157089/c
LOCUS       AX157089                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 417 from Patent WO0140521.
ACCESSION   AX157089
VERSION     AX157089.1  GI:14538420
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 417 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (418 is other entry)
                     Accession number cg44924574"
BASE COUNT       15 a      15 c      14 g      7 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 51;
  Best Local Similarity   67.4%;  Pred. No. 3.6e+06;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    1024 GCTTCTCGTCTGGGAGCTGGTCCTGGGGACGCTGGAGAAGTTC 1066
           |||||||  ||| |||||   |  |  |  ||||||  |||||
Db      48 GCTTCTCTCCTGAGAGCTCTCCGAGCTGGGGCTGGATCAGTTC 6


RESULT 85
AX157090/c
LOCUS       AX157090                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 418 from Patent WO0140521.
ACCESSION   AX157090
VERSION     AX157090.1  GI:14538421
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 418 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
```

```
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (417 is other entry)
                     Accession number cg44924574"
BASE COUNT       15 a     14 c     14 g      8 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 51;
  Best Local Similarity   67.4%;  Pred. No. 3.6e+06;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1024 GCTTCTCGTCTGGGAGCTGGTCCTGGGGACGCTGGAGAAGTTC 1066
           ||||||| ||| |||||   |  |  | ||||||  |||||
Db     48 GCTTCTCTCCTGAGAGCTCTCCAAGCTGGGGCTGGATCAGTTC 6


RESULT 86
AX158178
LOCUS       AX158178                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 1506 from Patent WO0140521.
ACCESSION   AX158178
VERSION     AX158178.1  GI:14539509
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 1506 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (1505 is other entry)
                     Accession number cg30144940"
BASE COUNT       15 a     19 c      9 g      8 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 51;
  Best Local Similarity   62.7%;  Pred. No. 3.6e+06;
  Matches   32;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy    669 CATCTACAGAGACCTGAAGCCGGAGAATATCATGCTTAATCACCAAGGTCA 719
          |||| |  | ||| | || |||   |||||| || | ||   ||  || ||
Db      1 CATCCATCGCGACGTCAAACCGACCAATATCTTGGTCAACACCCGCGGACA 51


RESULT 87
AX161887/c
LOCUS       AX161887                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 5215 from Patent WO0140521.
ACCESSION   AX161887
VERSION     AX161887.1  GI:14543218
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 5215 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (5216 is other entry)
                     Accession number cg43986974"
BASE COUNT       12 a     10 c     14 g     15 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 51;
  Best Local Similarity   67.4%;  Pred. No. 3.6e+06;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1145 TGCAATCTGAAGAGGATGTAAGTCAGTTTGATTCCAAGTTTAC 1187
           ||||   | ||| |||| |  | ||| || ||||| | || ||
Db     46 TGCAGAGTCAAGGGGATTTCCGACAGCTTCATTCCCAATTCAC 4


RESULT 88
AX161888/c
LOCUS       AX161888                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 5216 from Patent WO0140521.
ACCESSION   AX161888
VERSION     AX161888.1  GI:14543219
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 5216 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (5215 is other entry)
                     Accession number cg43986974"
BASE COUNT       13 a     10 c     13 g     15 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 51;
  Best Local Similarity   67.4%;  Pred. No. 3.6e+06;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1145 TGCAATCTGAAGAGGATGTAAGTCAGTTTGATTCCAAGTTTAC 1187
           ||||   | ||| |||| |  | ||| || ||||| | || ||
Db     46 TGCAGAGTCAAGGGGATTTCTGACAGCTTCATTCCCAATTCAC 4


RESULT 89
AX162091/c
LOCUS       AX162091                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 5419 from Patent WO0140521.
ACCESSION   AX162091
VERSION     AX162091.1  GI:14543422
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 5419 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
```

```
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (5420 is other entry)
                     Accession number cg43999946"
BASE COUNT       25 a      7 c      6 g     13 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 51;
  Best Local Similarity   74.3%;  Pred. No. 3.6e+06;
  Matches   26;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy    2098 ACAACCTGAATCTTTTTTTTATATAAATATATATT 2132
           |||||  | || ||||| ||  | |||||| ||||
Db      49 ACAACAGGTATGTTTTTGTTTCAGAAATATGTATT 15


RESULT 90
AX162092/c
LOCUS       AX162092                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 5420 from Patent WO0140521.
ACCESSION   AX162092
VERSION     AX162092.1  GI:14543423
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 5420 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (5419 is other entry)
                     Accession number cg43999946"
BASE COUNT       25 a      6 c      6 g     14 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 51;
  Best Local Similarity   74.3%;  Pred. No. 3.6e+06;
  Matches   26;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy    2098 ACAACCTGAATCTTTTTTTTATATAAATATATATT 2132
           |||||  | || ||||| ||  | |||||| ||||
Db      49 ACAACAGGTATGTTTTTGTTTCAAAAATATGTATT 15


RESULT 91
AX404671/c
LOCUS       AX404671                  53 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 45 from Patent WO0224745.
ACCESSION   AX404671
VERSION     AX404671.1  GI:21437952
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Abken,H. and Schinkoethe,T.
  TITLE     Method for detecting tumor cells
  JOURNAL   Patent: WO 0224745-A 45 28-MAR-2002;
            Abken, Hinrich (DE)
FEATURES             Location/Qualifiers
     source          1. .53
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       22 a      6 c      6 g     19 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 53;
  Best Local Similarity   62.7%;  Pred. No. 3.7e+06;
  Matches   32;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy    2109 CTTTTTTTTATATAAATATATATTTTTCAAATAGATTTTTGATTCAGCTCA 2159
           |||||| || ||  | |  | |  || | ||| |||| ||||  ||   ||
Db      52 CTTTTTCTTTTACTATTGAAAAAGTTGCTAATTGATTATTGAAACAATGCA 2


RESULT 92
AF328254/c
LOCUS       AF328254                  54 bp    DNA     linear   ROD 29-JUN-2001
DEFINITION  Mus musculus isolate 1.2-N1 T-cell receptor beta chain VDJ
            junctional region gene, partial cds.
ACCESSION   AF328254
VERSION     AF328254.1  GI:13898488
KEYWORDS    .
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Maryanski,J.L., Attuil,V., Hamrouni,A., Mutin,M., Rossi,M.,
            Aublin,A. and Bucher,P.
  TITLE     Individuality of Ag-selected and preimmune TCR repertoires
  JOURNAL   Immunol. Res. 23 (1), 75-84 (2001)
  MEDLINE   21310440
   PUBMED   11417861
REFERENCE   2  (bases 1 to 54)
  AUTHORS   Maryanski,J.L., Attuil,V., Hamrouni,A., Mutin,M., Rossi,M.,
            Aublin,A. and Bucher,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-DEC-2000) INSERM Unit 503, CERVI, 21 Avenue Tony
            Garnier, 69365 Lyon Cedex 07, France
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="Mus musculus"
                     /strain="DBA/2"
                     /isolate="1.2-N1"
                     /db_xref="taxon:10090"
                     /rearranged
     mRNA            <1. .>54
                     /product="T-cell receptor beta chain VDJ junctional
                     region"
     CDS             <1. .>54
                     /note="TCRBV10"
                     /codon_start=1
                     /product="T-cell receptor beta chain VDJ junctional
                     region"
                     /protein_id="AAK48763.1"
                     /db_xref="GI:13898489"
                     /translation="SAVYLCASSLQNSDYTFG"
BASE COUNT       11 a     18 c     10 g     15 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 10;  Length 54;
  Best Local Similarity   67.4%;  Pred. No. 3.7e+06;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    1638 AAGGTGGAGAGGGAGATGTGTGAGCATCCTGCAAGGTGAAACA 1680
           |||||| ||  |||| | ||| |||  |  |||  | || |||
Db      50 AAGGTGTAGTCGGAGTTTTGTAAGCTGCTGGCACAGAGATACA 8


RESULT 93
E10890
LOCUS       E10890                    60 bp    DNA     linear   PAT 29-SEP-1997
DEFINITION  PCR primer for gaining human cytochrome P450.
ACCESSION   E10890
VERSION     E10890.1  GI:22027985
KEYWORDS    JP 1996056695-A/39.
```

```
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Hayashi,K., Sakaki,T., Yabusaki,Y., Komai,K., Kaneko,H. and
            Nakatsuka,I.
  TITLE     METHOD FOR EVALUATING SAFETY
  JOURNAL   Patent: JP 1996056695-A 39 05-MAR-1996;
            SUMITOMO CHEM CO LTD
COMMENT     OS   None
            OC   Artificial sequences.
            PN   JP 1996056695-A/39
            PD   05-MAR-1996
            PF   15-JUL-1994 JP 1994164184
            PR   20-JUL-1993 JP 93P   201120, 30-JUL-1993 JP 93P   208279, PR
            17-JUN-1994 JP 94P   136053
            PI   HAYASHI KOJI, SAKAKI TOSHIYUKI, YABUSAKI YOSHIYASU, PI   KOMAI
            KOICHIRO,
            PI   KANEKO HIDEO, NAKATSUKA IWAO
            PC   C12Q1/02,C12M1/34,C12Q1/26;
            CC   strandedness: Single;
            CC   topology: Linear;
            CC   hypothetical: No;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .60
            FT                   /organism='Artificial sequences'.
FEATURES              Location/Qualifiers
     source           1. .60
                      /organism="unidentified"
                      /db_xref="taxon:32644"
BASE COUNT       24 a     12 c     17 g      7 t
ORIGIN

  Query Match             0.9%;  Score 20.6;  DB 6;  Length 60;
  Best Local Similarity   62.7%;  Pred. No. 3.7e+06;
  Matches   32;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy    888 GACTGGAGCACCCCCATTCACTGGGGAGAATAGAAAGAAAACAATTGACAA 938
          || || |||||  |   |    |||  | | ||||  ||| | ||||||||
Db     10 GAATGAAGCACAGCTGATAGAAGGGCGGGACAGAAGCAAATCCATTGACAA 60


RESULT 94
AX117117
LOCUS       AX117117                 51 bp    DNA     linear   PAT 11-MAY-2001
DEFINITION  Sequence 2240 from Patent WO0129262.
ACCESSION   AX117117
VERSION     AX117117.1  GI:14034059
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Picoult-Newburg,L. and Pohl,M.
  TITLE     Genotyping reagents, kits and methods of use thereof
  JOURNAL   Patent: WO 0129262-A 2240 26-APR-2001;
            Orchid BioSciences, Inc. (US)
FEATURES              Location/Qualifiers
     source           1. .51
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
BASE COUNT       25 a      7 c      7 g     11 t      1 others
ORIGIN

  Query Match             0.9%;  Score 20.4;  DB 6;  Length 51;
  Best Local Similarity   65.2%;  Pred. No. 4.1e+06;
  Matches   30;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   1724 AACACTTCGGACACAGGAAAAATAAACGTGGATTTTAAAAAATCAA 1769
          | || |||  ||| ||| |||| ||   |  |||  | ||||||
Db      2 AGCAATTCAAACAGAGGGAAAAAAACTATAAATTGCAGAAAATCTA 47
```

```
RESULT 95
YSCMTP182
LOCUS       YSCMTP182                57 bp    DNA     linear   PLN 04-AUG-1993
DEFINITION  Yeast (S.cerevisiae) mitochondrial petite mutant excision seq 18,
            right end.
ACCESSION   J01505
VERSION     J01505.1  GI:343883
KEYWORDS    AT-rich region; GC rich region.
SEGMENT     2 of 2
SOURCE      Yeast (Saccharomyces cerevisiae) mitochondrial DNA.
  ORGANISM  Mitochondrion Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 57)
  AUTHORS   de Zamaroczy,M., Faugeron-Fonty,G. and Bernardi,G.
  TITLE     Excision sequences in the mitochondrial genome of yeast
  JOURNAL   Gene 21 (3), 193-202 (1983)
  MEDLINE   83210931
   PUBMED   6343188
COMMENT     Additional sequences reported in [1], but sequenced in earlier
            papers, appear in separate entries. Excision repeat corresponds to
            bases 11 to 25.
FEATURES              Location/Qualifiers
     source           1. .57
                      /organism="Saccharomyces cerevisiae"
                      /organelle="mitochondrion"
                      /db_xref="taxon:4932"
BASE COUNT       23 a      1 c      1 g     31 t      1 others
ORIGIN

  Query Match             0.9%;  Score 20.4;  DB 8;  Length 57;
  Best Local Similarity   60.0%;  Pred. No. 4.1e+06;
  Matches   33;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy   2118 ATATAAATATATATTTTTCAAATAGATTTTTGATTCAGCTCATTATGAAAAACAT 2172
          ||| ||   | ||||| | |  ||| |||||  |      |||| | | ||| ||
Db      2 ATAAAATATTTTATTTATTATTTAGNTTTTTTTTATTAAACATTTTTATAAAAAT 56


RESULT 96
A08918
LOCUS       A08918                   39 bp    DNA     linear   PAT 02-SEP-1993
DEFINITION  H.sapiens (allele MS32, isolate English, serial number 3 and 4)
            minisatellite sequence.
ACCESSION   A08918
VERSION     A08918.1  GI:411840
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Jeffreys,A.J.
  TITLE     Extended nucleotide sequences
  JOURNAL   Patent: EP 0370719-A 101 30-MAY-1990;
            IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES              Location/Qualifiers
     source           1. .39
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
BASE COUNT       18 a      0 c      0 g     21 t
ORIGIN

  Query Match             0.9%;  Score 20.2;  DB 6;  Length 39;
  Best Local Similarity   75.8%;  Pred. No. 4.4e+06;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   2108 TCTTTTTTTTATATAAATATATATTTTTCAAAT 2140
          | | |||||||||| ||| | | |||| | ||||
Db      7 TATATTTTTTATAAAAAAAAAAATTTATTAAAT 39
```

```
RESULT 97
HSU26975/c
LOCUS       HSU26975                  45 bp    mRNA    linear   PRI 10-JUN-1995
DEFINITION  Human isolate M15 T-cell receptor V-alpha 3/J alpha 57 junction
            mRNA, partial cds.
ACCESSION   U26975
VERSION     U26975.1  GI:857421
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Dave,V.P., Larche,M., Rencher,S.D., Koop,B.F. and Hurwitz,J.L.
  TITLE     Restricted usage of T-cell receptor V alpha sequence and
            variable-joining pairs after normal T-cell development and bone
            marrow transplantation
  JOURNAL   Hum. Immunol. 37 (3), 178-184 (1993)
  MEDLINE   94064390
   PUBMED   8244780
REFERENCE   2  (bases 1 to 45)
  AUTHORS   Hurwitz,J.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-1995) Julia L. Hurwitz, Immunology, St. Jude
            Children's Research Hospital, 332 N. Lauderdale, Memphis, TN 38101,
            USA
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="Homo sapiens"
                     /isolate="M15"
                     /db_xref="taxon:9606"
                     /tissue_type="blood"
     CDS             <1. .>45
                     /note="encodes V alpha 3/J alpha 57 junction"
                     /codon_start=1
                     /evidence=experimental
                     /product="T-cell receptor"
                     /protein_id="AAA68158.1"
                     /db_xref="GI:857422"
                     /translation="FCATDPWGGSEKLVF"
BASE COUNT       7 a      10 c      14 g      14 t
ORIGIN

  Query Match             0.9%;  Score 20.2;  DB 9;  Length 45;
  Best Local Similarity   68.3%;  Pred. No. 4.5e+06;
  Matches   28;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   1547 AAAAACAAGCTTTTCCCATGATCTCCAAACGGCCAGAGCAC 1587
          ||| || ||||||||  ||   | ||||  | ||  |||||
Db     45 AAAGACCAGCTTTTCAGATCCGCCCCAAGGGTCCGTAGCAC 5


RESULT 98
AR032407
LOCUS       AR032407                  48 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 19 from patent US 5869241.
ACCESSION   AR032407
VERSION     AR032407.1  GI:5948012
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Method of determining DNA sequence preference of a DNA-binding
            molecule
  JOURNAL   Patent: US 5869241-A 19 09-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="unknown"
BASE COUNT       20 a       8 c       6 g      14 t
ORIGIN

  Query Match             0.9%;  Score 20.2;  DB 6;  Length 48;
  Best Local Similarity   68.3%;  Pred. No. 4.5e+06;
  Matches   28;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   2082 GATGTTTTACGTGCAAACAACCTGAATCTTTTTTTTATATA 2122
          ||||||  ||  ||||||||  | ||| | | |   |||||
Db      1 GATGTTACACAAGCAAACAAAATAAATATCTGTGCAATATA 41


RESULT 99
AR209071
LOCUS       AR209071                  48 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 19 from patent US 6384208.
ACCESSION   AR209071
VERSION     AR209071.1  GI:21510392
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Sequence directed DNA binding molecules compositions and methods
  JOURNAL   Patent: US 6384208-A 19 07-MAY-2002;
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="unknown"
BASE COUNT       20 a       8 c       6 g      14 t
ORIGIN

  Query Match             0.9%;  Score 20.2;  DB 6;  Length 48;
  Best Local Similarity   68.3%;  Pred. No. 4.5e+06;
  Matches   28;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   2082 GATGTTTTACGTGCAAACAACCTGAATCTTTTTTTTATATA 2122
          ||||||  ||  ||||||||  | ||| | | |   |||||
Db      1 GATGTTACACAAGCAAACAAAATAAATATCTGTGCAATATA 41


RESULT 100
I29147
LOCUS       I29147                    48 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 19 from patent US 5578444.
ACCESSION   I29147
VERSION     I29147.1  GI:1819938
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Sequence-directed DNA-binding molecules compositions and methods
  JOURNAL   Patent: US 5578444-A 19 26-NOV-1996;
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="unknown"
BASE COUNT       20 a       8 c       6 g      14 t
ORIGIN

  Query Match             0.9%;  Score 20.2;  DB 6;  Length 48;
  Best Local Similarity   68.3%;  Pred. No. 4.5e+06;
  Matches   28;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   2082 GATGTTTTACGTGCAAACAACCTGAATCTTTTTTTTATATA 2122
          ||||||  ||  ||||||||  | ||| | | |   |||||
Db      1 GATGTTACACAAGCAAACAAAATAAATATCTGTGCAATATA 41


Search completed: April 19, 2003, 08:53:58
Job time : 6855 secs
```